

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 21, 2004, 20:33:12 ; Search time 530 Seconds  
(without alignments)  
2605.027 Million cell updates/sec

Title: US-09-931-457a-31  
Perfect score: 1623  
Sequence: 1 MAVERSGIAKDVTEIGKTP.....LSVLFESVREAESMTPEP 325

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DBV=x1b  
-Q=/cgn2.1/USPTO\_spool/US09931457/runat\_18052004\_121728\_20289/app\_query.fasta\_1.519  
-DB=N Geneseg 29Jan04 -QFMT=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09931457 @CGN 1.1 470 @runat\_18052004\_121728\_20289 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseg 29Jan04: \*  
1: geneseg1980s: \*  
2: geneseg1990s: \*  
3: geneseg2000s: \*  
4: geneseg2001as: \*  
5: geneseg2001bs: \*  
6: geneseg2002s: \*  
7: geneseg2003as: \*  
8: geneseg2003bs: \*  
9: geneseg2003cs: \*  
10: geneseg2004as: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1623	100.0	1362	2	AAV99906 Fragment
2	1431	88.2	1303	2	AAQ44450 Cysteine
3	1388	85.5	946	9	ADC76317 DNA homol
4	1384	85.3	1120	3	AAC42360 Arabidops
5	1374	84.7	966	6	ABZ13088 Arabidops
6	1374	84.7	966	7	ADA68423 Arabidops
7	1360	83.8	993	7	ADA69412 Rice gene
8	1359	83.7	5259	7	ADA70437 Rice gene

9	1271	78.3	1483	2	AAQ74413 Spinach c
10	1261	77.7	1503	3	AAC34662 Arabidops
11	1255	77.3	1329	6	ABZ12294 Arabidops
12	1196	73.7	1224	3	AAC40774 Arabidops
13	1192	73.4	972	6	ABZ12690 Arabidops
14	1192	73.4	999	3	AAC45719 Arabidops
15	1192	73.4	1423	3	AAC41206 Arabidops
16	1191	73.4	1148	3	AAC47965 Arabidops
17	1184	73.0	1303	3	AAC33559 Arabidops
18	1179	72.6	975	9	ADC64221 Thalecres
19	1179	72.6	1200	3	AAC44727 Arabidops
20	1166	71.8	1302	3	AAC45323 Arabidops
21	1149	70.8	1215	3	AAC44893 Arabidops
22	1129	69.6	1037	9	ADC76318 DNA homol
23	1125	69.3	1006	9	ADC75697 DNA homol
24	1092	67.3	1030	9	ADD17590 DNA (Seq)
25	1077	66.4	725	9	ADD17591 DNA (Seq)
26	1065.5	65.7	918	3	AAC41635 Arabidops
27	1045	64.4	1481	3	AAC47004 Arabidops
28	1045	64.4	1529	4	ABA01246 O-acetyl
29	1017	62.7	1431	3	AAC46837 Arabidops
30	1006.5	62.0	1417	3	AAC51477 Arabidops
31	982	60.5	1089	3	AAC51458 Arabidops
32	980.5	60.4	683	9	ADC75696 DNA homol
33	974.5	60.0	1200	7	ADA69647 Rice gene
34	973	60.0	657	8	ACL17131 DNA clone
35	968.5	59.7	945	7	AAC24366 Prokaryot
36	962	59.3	913	9	ADC76310 DNA homol
37	962	59.3	933	9	ADD16769 DNA (Seq)
38	949.5	58.5	930	7	AAC38650 Prokaryot
39	949.5	58.5	933	7	ACA40638 Prokaryot
40	949.5	58.5	110000	4	Continuation (26 c
41	949.5	58.5	110000	4	Continuation (27 c
42	949.5	58.5	110000	4	Continuation (26 c
43	949.5	58.5	110000	4	Continuation (27 c
44	941.5	58.0	930	7	ACA38024 Prokaryot
45	941.5	58.0	1086	7	ACA38210 Prokaryot

## ALIGNMENTS

RESULT 1  
AAV99906  
ID AAV99906 standard; cDNA; 1362 BP.

XX AC AAV99906;

DT 27-SBP-1999 (first entry)

XX Fragment of cysteine synthase gene.

XX Biosynthesis; biosynthetic pathway; lysine; threonine; methionine;

KW cysteine; isoleucine; amino acid; homoserine kinase;

KW aspartic semialdehyde dehydrogenase; diaminopimelate decarboxylase;

KW cysteine synthase; cystathione beta-lyase; gene expression; screening;

XX inhibition; ss.

OS Glycine max.

XX Key

PN W09856935-A2.

XX 17-DEC-1998.

XX 11-JUN-1998; 98WD-US012073.

XX 12-JUN-1997; 97US-0049406P.

XX 12-NOV-1997; 97US-0065385P.

Location/Qualifiers  
90.1067  
/\*tag= a  
/product= "Cysteine synthase fragment"

PA (DUPLO ) DU PONT DE NEMOURS & CO E I.  
 XX Falco SC, Allen SM, Thorpe CJ;  
 XX WPI; 1999-080910/07.  
 DR P-FSDB; AAW81018.  
 XX  
 PT New isolated plant amino acid biosynthetic enzyme nucleic acids - which  
 PT encode aspartic semialdehyde dehydrogenase, diaminopimelate  
 PT decarboxylase, homoserine kinase, cysteine synthase and cystathionine  
 PT beta-lyase.  
 XX  
 PS Claim 17; Page 57-58; 80pp; English.

CC Organisation of the pathway leading to plant biosynthesis of lysine,  
 CC threonine, methionine, cysteine and isoleucine suggests that over-  
 CC expression or reduction of expression of genes encoding enzymes involved  
 CC in that biosynthetic pathway could be used to alter the level of these  
 CC amino acids in human food and animal feed. This may increase the  
 CC nutritional quality of human food and animal feed by increasing the  
 CC production and accumulation of specific free amino acids. The enzymes  
 CC include aspartic semialdehyde dehydrogenase, homoserine kinase,  
 CC diaminopimelate decarboxylase, cysteine synthase and cystathione beta-  
 CC lyase. The nucleic acids encoding these enzymes can be used for altering  
 CC the level of expression of the enzymes and for evaluating compounds for  
 CC their ability to inhibit the enzymes' activity

XX Sequence 1362 BP; 388 A; 243 C; 330 G; 401 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 3.25e-151 Length: 1362  
 Score: 1623.00 Matches: 325  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-09-931-457A-31 (1-325) x AAV99906 (1-1362)

QY 1 MetAlaValGluArgSerGlyIleAlaIysAspValThrGluLeuIleGlyLysThrPro 20  
 DB 90 ATGGCTGTGAAGGTCGCCGAATGCGCAAGATGTACGGAATGATGTGTAACCCCA 149  
 QY 21 LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40  
 DB 150 TTAGTATATCTAATAAATTAACCTTGGGATGTGTGTCGCCGGTGTGCTGCTAACTGAG 209  
 QY 41 LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60  
 DB 210 TTGATGAGGCCATCTCTAGTGTGAAGACAGCATGGGTATAGTATGATGCTGATGCA 269  
 QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80  
 DB 270 GAAGAGAGGAGCTTATCAACCTGGAAGAGTGTCTCTCATTTGAGCCCAACAGTGAAT 329  
 QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaArgGlyTyrLysLeuIleIleThrMet 100  
 DB 330 ACTGGCATTTGATAGCTTCTTATGCGAGCAGCCAGGGTTACAAAGCTCATTAATTAATG 389  
 QY 101 ProAlaSerMetSerLeuGluArgGlyIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120  
 DB 390 CCGCTTCTATAGAGTCTTGAAGAGAAATCATTTCTATTAGCTTTTGGAGCTGAGTGGTT 449  
 QY 121 LeuThrAspProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluGluIleLeuAla 140  
 DB 450 CTGACAGATCTCTGCTAAGGGAATGAAGAGTGTCTTCAAGAGGCTGAAGAGATATTGGCT 509  
 QY 141 LysThrProAsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHis 160  
 DB 510 AAGAGCGCCCAATGCTTACATACTTCAACAATTTGAAGAAACCTGCGCAATCCCAAGGTTCT 569  
 QY 161 TyrGluThrThrGlyProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheVal 180

DB 570 TATGAACCACTGCTCCAGAGATATGGAAGGCTCCGATGGGAAATTGATGCAATTGTT 629  
 QY 181 SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsn 200  
 DB 630 TCTGGGATAGGCACTGGTGATCAATTAACAGGTGCTGGAAAAATATCTTAAGAGAGCAAT 689  
 QY 201 ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLys 220  
 DB 690 CCGATATTAAGCTGATGTGTGGAACCACTGAAAGTCCAGTGTCTCAAGAGAGAAAG 749  
 QY 221 ProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240  
 DB 750 CCGTGTCCACACAAAGATTCAAGGAGATTGTGTGTGTTTATCCCTGTGTCTTGAAGTC 809  
 QY 241 AsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260  
 DB 810 AATCTTCTTGATGAAGTGTTCAAATATCAAGTGAATGACCAATAGAACTGCAAGCTT 869  
 QY 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla 280  
 DB 870 CTTCGGCTTAAGAGGCTTAATTTGTGGAAATATCTCCGAGCTGCACTGTGCTGCT 929  
 QY 281 PheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSer 300  
 DB 930 TTTCAGATTGCAAAAGACCAAAATATGCGGAGAGCTTAATGTGCTGCTTTTCCAGC 989  
 QY 301 PheGlyGluArgTyrLeuSerSerValLeuPheGluSerValArgGluValGluSer 320  
 DB 990 TTCGGGAGAGAGTACTGTCTCCGTGCTATTGAGTCACTGAGACGCGAAGCTGAAGC 1049  
 QY 321 MetThrPheGluPro 325  
 DB 1050 ATGACTTTGAGCCC 1064  
 RESULT 2  
 AAQ44450  
 ID AAQ44450 standard; cDNA to mRNA; 1303 BP.  
 XX AAQ44450;  
 AC 12-OCT-1994 (first entry)  
 DT 12-OCT-1994 (first entry)  
 XX  
 DE Cysteine synthase.  
 KW Cysteine synthase; plant; expression; probe; ss.  
 XX Spinacia oleracea.  
 FH Key Location/Qualifiers  
 FT CDS 52..1029  
 FT /\*tag= a  
 XX JP06038770-A.  
 PN 15-FEB-1994.  
 XX PD  
 XX 05-FEB-1992; 92JP-00020315.  
 PF 05-FEB-1992; 92JP-00020315.  
 XX PR 05-FEB-1992; 92JP-00020315.  
 XX  
 PA (MITS ) MITSUBISHI CORP.  
 PA (MITU ) MITSUBISHI KASEI CORP.  
 XX  
 DR WPI; 1994-094834/12.  
 DR P-FSDB; AAR49830.  
 XX  
 PT Novel gene coding cysteine synthase - used to increase the cysteine  
 PT content of an agricultural plant.  
 PS Claim 2; Page 4-6; 6pp; Japanese.  
 CC The cysteine content in an agricultural product can be increased by  
 CC expressing the cysteine synthase in a plant. Probes V822 and V812 used in

CC the isolation of the gene are given in AAQ44483-84  
XX Sequence 1303 BP, 367 A, 242 C, 311 G, 383 T, 0 U, 0 Other;

Alignment Scores:

Pred. No.:	3.38e-132	Length:	1303
Score:	1431.00	Matches:	281
Percent Similarity:	94.14%	Conservative:	24
Best Local Similarity:	86.73%	Mismatches:	19
Query Match:	88.17%	Indels:	0
DB:	2	Gaps:	0

US-09-931-457A-31 (1-325) x AAQ44450 (1-1303)

QY 1 MetAlaValGluArgSerGlyIleAlaIysAspValThrGluLeuIleGlyLysThrPro 20  
DB 52 ATGGTTGAGGAGAGGCTTCATGCTAAAGATGTAATGATTGATGGGAAACGCCA 111  
QY 21 LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaIleLysLeuGlu 40  
DB 112 TTGGTATATCTCAACACTGCGCCGATGTTGTTGCTCGTGTGCTGCAAGCTGGA 171  
QY 41 LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60  
DB 172 GGAATGGAACCTTGCTCTAGTGTAAAGACAGAGATTGGTTCACTATGATTACTGATGCT 231  
QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80  
DB 232 GAAAAAAGCGGCTTATACCTCGAGAGAGTGTCTGATTGAGCCACGAGGAAAT 291  
QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMet 100  
DB 292 ACTGGCAATTGATTAGCTTCACTGACAGCACTAAAGCTTACAGCTCATTCATTCATGATG 351  
QY 101 ProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120  
DB 352 CCAGCATCAATAGTCTTGAGCGGAGAGACTATTCTCAGGGCTTTGTTGCTGAGCTTATC 411  
QY 121 LeuThrAspProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluGluIleLeuAla 140  
DB 412 CTTAAGTATCCAGCAAAAGGTATGAAGGGCTGTTCAAGAGGCTGAGAGATCCGTGAC 471  
QY 141 LysThrProAsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHis 160  
DB 472 AAAAATCTTAATTCATATATTAACAACAGTTGAAAACCTGCCAACCCAAAGCTTCAT 531  
QY 161 TyrGluThrThrGlyProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheVal 180  
DB 532 TATGAAACAACACTGACCAAGAAATTTGAAAGGACACAGCTGGAATAAATTGATATTCGTC 591  
QY 181 SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsn 200  
DB 592 TCTGGAATAGGAGACTGAGGTACAATAACAGGTGACGAAATACTTAAAGAAACAAAC 651  
QY 201 ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLys 220  
DB 652 CCGGATGTTAAGCTAATGCGCTGGAACCAAGTGAAGTCTGTATTGCTGAGGAAAA 711  
QY 221 ProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240  
DB 712 CCGGCCCAACATAAGATTCAAGACTTGAGCTGATTCATACCTGCTGCTGCGATGTCG 771  
QY 241 AsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260  
DB 772 AATATATATGATGAGTGTGATATCCAGTGAAGATCTAATTGAATGCGCAATG 831  
QY 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla 280  
DB 832 CTCGCCCTCAAGGAGCTCTACTGCTTGGGATTTCATCTGCTGCTGCTGCGCTGCC 891  
QY 281 PheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSer 300  
DB 892 AATTAAGTGGCAAGAGGCTCAAAATGCTGGAATACTCATCTGCTGCTCTTCCACG 951

QY 301 PheGlyGluArgTyrLeuSerSerValLeuPheGluSerValArgGluAlaGluSer 320  
DB 952 TTGGCGAAGCATATTATTCCTCGGTGTGTTGATTCACTGAGGAGAGAGAGAGC 1011  
QY 321 MetThrPheGlu 324  
DB 1012 ATGGTATTGAG 1023

RESULT 3

ADCT6317  
ID ADCT6317 standard; DNA; 946 BP.  
XX  
AC ADCT6317;  
XX  
DT 01-JAN-2004 (first entry)  
XX  
DE DNA homologous to phytopathogen resistance-related cDNA - SEQ ID 1586.  
XX  
KW rice; yeast; poppy; plant; disease resistance; anti-fungal;  
XX  
KM phytopathogen; gene shuffling; ds.  
XX  
OS Unidentified.  
XX  
PN W02003020905-A2.  
XX  
PD 13-MAR-2003.  
XX  
PF 30-AUG-2002; 2002WO-US027883.  
XX  
PR 31-AUG-2001; 2001US-0316392P.  
XX  
PA (DOWC) DOW CHEM CO.  
XX  
PI Shukla V, Butler H, Larrinua I, Reddy AS;  
XX  
DR WPI; 2003-290185/28.  
XX

PT Novel isolated nucleic acid derived from *Nicotiana benthamiana*, *Oryza*  
PT *sativa*, *Saccharomyces cerevisiae*, *Trichoderma harzianum* and *Papaver*  
PT *rheas*, useful for conferring disease resistance in plants.  
XX

PS Claim 1: SEQ ID NO 1586; 617bp; English.

XX The invention relates to a novel isolated nucleic acid derived from  
CC *Nicotiana benthamiana*, *Oryza sativa* (rice), *Saccharomyces cerevisiae*  
CC (yeast), *Trichoderma harzianum* (Hypocrea lixi) and *Papaver rheas*  
CC (poppy) and a sequence that hybridises to them under conditions of low  
CC stringency, where expression of the nucleic acid in a plant results in a  
CC disease resistance phenotype. The polynucleotides of the invention  
CC demonstrate anti-fungal activity and may be useful in conferring disease  
CC resistance in a plant against phytopathogen such as *Aspergillus flavus*,  
CC *Gibberella fujikuroi* and *Gibberella zeae*. Furthermore, the  
CC polynucleotides may be useful to retrieve unknown sequences and in gene  
CC shuffling or sexual PCR procedures. The current sequence is that of the  
CC DNA of the invention which is homologous to that of the phytopathogen  
CC resistance-related config cDNAs.

XX Sequence 946 BP; 259 A; 176 C; 250 G; 261 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.07e-128	Length:	946
Score:	1388.00	Matches:	269
Percent Similarity:	94.21%	Conservative:	24
Best Local Similarity:	86.50%	Mismatches:	18
Query Match:	85.52%	Indels:	0
DB:	9	Gaps:	0

US-09-931-457A-31 (1-325) x ADCT6317 (1-946)

QY 2 AlavaGluArgSerGlyIleAlaIysAspValThrGluLeuIleGlyLysThrProIeu 21  
DB 13 GCGGAGAGAAAGATGGAATTGCGAAGATGTAATGATTGATCGGTAAACACTCTTGG 72

QY 22 ValTyrLeuasnLysLeuAlaaspGlyCysValAlaargValAlaAlaLysLeuGluLeu 41  
DB 73 GTGACCTGATTAATGTTGTGATGGGTGTTGCTCGCGTTGCTGCAAGCTCGAAAGC 132  
QY 42 MetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaaspAlaGlu 61  
DB 133 ATGAGCCATGCTCTAGTGTAAAGATAGATGCTTATAGTATGATTAACAGATGCTGAG 192  
QY 62 GluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsnThr 81  
DB 193 GAGAAATGGCTGATCAAACTGGCGAGAGTGTCTCATTTGAACCTACAAAGTGAACACT 252  
QY 82 GlyIleGlyLeuAlaPheMetAlaAlaAlaargGlyTyrLysLeuIleIleThrMetPro 101  
DB 253 GGAGTAGATGGCATTTATGGCTGTCTGCTAAAGGCTACAAACTCATTAACGATGCT 312  
QY 102 AlaSerMetSerLeuGluArgGlyIleIleLeuLeuAlaPheGlyAlaGluLeuValLeu 121  
DB 313 TCTCAATGAGTCTTGAGAGAGAAATTTTGGCTGCTTTGGTGTGATGCTGCTGCT 372  
QY 122 ThrAspProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluGlnIleLeuAlaLys 141  
DB 373 ACCGATCCAGCAAAAGGATGAAGGTTCTATTCAGAAGGCTGAAGGATTAAGGCCAAA 432  
QY 142 ThrProAspAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHistyr 161  
DB 433 ACACCTAACTCTTATTTCTTACGCAATTGAAGAAACCTGCAAAACCAAGTACACTAT 492  
QY 162 GluThrThrGlyProGluIleTyrPheGlySerAspGlyLysIleAspAlaPheValSer 181  
DB 493 GAGACAACTGCTCTGATGATCTGGAAGGCTCAAAACGGAAGTACTGCTCTCTCT 552  
QY 182 GlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsnPro 201  
DB 553 GGAATTGAAACAGAGGCAACAATACAGGTTCAAGCAAGTATTTAAGAGACAGAACCCC 612  
QY 202 AsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLysPro 221  
DB 613 GACATAAAGCTCTATGCTGTGGAACCAAGTGAAGTCTATTCTTCTGAGGAATCCT 672  
QY 222 GlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsn 241  
DB 673 GGTCCGATTAAGTTCAGGGGATGTTGCTGCTTCTCTGCTGCTGCTGCTGCTGCT 732  
QY 242 LeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeuLeu 261  
DB 733 CTATTGATGAAGTAATCAAGTTTCAAGTGAAGAACCATAGAACTGCAAGCTTCTG 792  
QY 262 AlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAlaPhe 281  
DB 793 GCATTAAAGGAAGGTTGCTTGTGCGCAATATCTGCTGCTTCTGCTGCTGCAACATC 852  
QY 282 GlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSerPhe 301  
DB 853 AAACCTTGCTAAGCGCCCTGAAATGCTGGAGCTGATGTTGTTTCCCAAGCTTC 912  
QY 302 GlyGluArgTyrLeuSerSerValLeuPheGlu 312  
DB 913 GGGAGCGATATCTTCTCTGCTGCTTGA 945

## RESULT 4

AAC42360  
ID AAC42360 standard; DNA; 1120 BP.

AC AAC42360;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 35256.

XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;

KW promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
XX EP1033405-A2.  
XX 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
XX 05-MAR-1999; 99US-0123180P.  
XX 09-MAR-1999; 99US-0123548P.  
XX 23-MAR-1999; 99US-0125788P.  
XX 25-MAR-1999; 99US-0126264P.  
XX 29-MAR-1999; 99US-0126785P.  
XX 01-APR-1999; 99US-0127462P.  
XX 06-APR-1999; 99US-0128234P.  
XX 08-APR-1999; 99US-0128714P.  
XX 16-APR-1999; 99US-0129845P.  
XX 19-APR-1999; 99US-0130077P.  
XX 21-APR-1999; 99US-0130449P.  
XX 23-APR-1999; 99US-0130510P.  
XX 28-APR-1999; 99US-0130891P.  
XX 30-APR-1999; 99US-0131449P.  
XX 04-MAY-1999; 99US-0132048P.  
XX 05-MAY-1999; 99US-0132407P.  
XX 06-MAY-1999; 99US-0132484P.  
XX 07-MAY-1999; 99US-0132485P.  
XX 11-MAY-1999; 99US-0132486P.  
XX 14-MAY-1999; 99US-0132487P.  
XX 14-MAY-1999; 99US-0132863P.  
XX 14-MAY-1999; 99US-0134218P.  
XX 14-MAY-1999; 99US-0134219P.  
XX 14-MAY-1999; 99US-0134221P.  
XX 14-MAY-1999; 99US-0134370P.  
XX 18-MAY-1999; 99US-0134768P.  
XX 19-MAY-1999; 99US-0134941P.  
XX 20-MAY-1999; 99US-0135124P.  
XX 21-MAY-1999; 99US-0135353P.  
XX 24-MAY-1999; 99US-0135629P.  
XX 25-MAY-1999; 99US-0136021P.  
XX 27-MAY-1999; 99US-0136392P.  
XX 28-MAY-1999; 99US-0136782P.  
XX 01-JUN-1999; 99US-0137222P.  
XX 03-JUN-1999; 99US-0137528P.  
XX 04-JUN-1999; 99US-0137502P.  
XX 07-JUN-1999; 99US-0137724P.  
XX 08-JUN-1999; 99US-0138094P.  
XX 10-JUN-1999; 99US-0138540P.  
XX 10-JUN-1999; 99US-0138847P.  
XX 14-JUN-1999; 99US-0139119P.  
XX 16-JUN-1999; 99US-0139452P.  
XX 16-JUN-1999; 99US-0139453P.  
XX 17-JUN-1999; 99US-0139492P.  
XX 18-JUN-1999; 99US-0139454P.  
XX 18-JUN-1999; 99US-0139455P.  
XX 18-JUN-1999; 99US-0139456P.  
XX 18-JUN-1999; 99US-0139457P.  
XX 18-JUN-1999; 99US-0139458P.  
XX 18-JUN-1999; 99US-0139459P.  
XX 18-JUN-1999; 99US-0139460P.  
XX 18-JUN-1999; 99US-0139461P.  
XX 18-JUN-1999; 99US-0139462P.  
XX 18-JUN-1999; 99US-0139463P.  
XX 18-JUN-1999; 99US-0139750P.  
XX 18-JUN-1999; 99US-0139763P.  
XX 21-JUN-1999; 99US-0139817P.  
XX 22-JUN-1999; 99US-0139899P.  
XX 23-JUN-1999; 99US-0140353P.  
XX 24-JUN-1999; 99US-0140354P.  
XX 24-JUN-1999; 99US-0140695P.



PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.

PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0158029P.  
PR 07-OCT-1999; 99US-0158232P.  
PR 08-OCT-1999; 99US-0158369P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:  
Pred. No.: 1,276-127  
Score: 1384.00  
Percent Similarity: 92.79%  
Best Local Similarity: 83.39%  
Query Match: 85.27%  
DB: 3  
Length: 1120  
Matches: 266  
Conservative: 30  
Mismatch: 23  
Indels: 0  
Gaps: 0

US-09-931-457A-31 (1-325) x AAC42360 (1-1120)

QY 6 SerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrProLeuValTyrLeuAsn 25  
|||  
Db 156 TCAGAGATTGCTAAGATGTGACTGAATTGATGGGAACACTCCATTGGTATTTGAAC 217  
QY 26 LysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluProCys 45  
:::|  
Db 218 AATGTGCTGAAGATGTGTGTCGTGCTGCTGAAGCTTGAGATGACAAACCGTGC 277  
QY 46 SerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGluLysGlyLeu 65  
|||  
Db 278 TCTAGCGTCAAGACAGATTTGTTTACATGATTTCTGATGCAGAAAGAGGCTCTT 337  
QY 66 IleThrProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIleGlyLeu 85  
|||  
Db 338 ATCAACACGAGACAGACTGTGCTGATTCAGCAACAAGTGAACACTGAGTGGGTTA 397  
QY 86 AlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProAlaSerMetSer 105  
|||

DB 398 GCATTCACGGCAGCTGCCAAGGCTACAGCTTATATACATGCCAGCTCTATGAGT 457  
QY 106 LeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspProAla 125  
DB 458 ACTGAGAGAGAAATCATCTCTTAGCTTTTGAGCTGAGTTGTTTAACTGACCCAGCT 517  
QY 126 LysGlyMetLysGlyAlaValGlnLysAlaGluGlnIleLeuAlaLysThrProAsnAla 145  
DB 518 AAGGCATGAAGAGAGCTATCGCAAGCGGAGAGATTTTGCGGAAAACACCCATGCT 577  
QY 146 TyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThrGly 165  
DB 578 TACATGCTTCAGCAGTTTGAGAACCCCTGCCAACCTTAAGATCCATAGACTACGGGA 637  
QY 166 ProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheValSerGlyIleGlyThr 185  
DB 638 CCTGAGATATGAAAGGCACTGCTGCGCAAAATCGATGCTTTGTTCTGCGATTGCTACT 697  
QY 186 GlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsnProAsnIleLysLeu 205  
DB 698 GGTGTACCATTAAGAGTCTGCGAAGTATCTTAAGACAGAACGCAACGCTCAAGCTG 757  
QY 206 IleGlyValGluProValGluSerProValLeuSerGlyLysProGlyProHisLys 225  
DB 758 TATGAGTGGAGCCAGTTGAAAGTCTATTCATTCCTGCTGGGAGCCAGCTCTCAACAG 817  
QY 226 IleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeuLeuAspGlu 245  
DB 818 ATTCAAGGATAGAGCTGTTTATATACCAAGTGTATGAATGATGATCTTATTGACGAA 877  
QY 246 ValValGlnIleSerSerAspGluAlaIleGluThrAlaLysIleLeuAlaLeuLysGlu 265  
DB 878 GTTGTTCAGGTTTCAGATGATGATCAATTCATTCATGCAAGGAGGAGCTTCTTAAAGAA 937  
QY 266 GlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAlaAlaPheGlnIleAlaLys 285  
DB 938 GGTCTTCTGTGGGATATCATCCGTCGACGAGCTGCTGACCAATTAACTTGACACAG 997  
QY 286 ArgProGluAsnAlaGlyLysIleuIleValAlaValPheProSerPheGlyGluArgTyr 305  
DB 998 AGGCCAGAAAACGCTGGGAGAGCTATTGTGGCGATATTCGCGAGTTTCGGGAGAGAGTAT 1057  
QY 306 LeuSerSerValLeuPheGluSerValArgArgGluAlaGluSerMetThrPheGlu 324  
DB 1058 CTATCAACGCTACTTTCGATGCGACAGAGAAGAGCGGAGCCATGACCTTCGAG 1114  
RESULT 5  
ABZ13088 standard; DNA; 966 BP.  
ID ABZ13088 standard; DNA; 966 BP.  
XX ABZ13088;  
AC 21-JAN-2003 (first entry)  
DT 21-JAN-2003 (first entry)  
XX Arabidopsis thaliana stress regulated gene SEQ ID NO 893.  
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 893.  
XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
KW Arabidopsis thaliana.  
OS Arabidopsis thaliana.  
XX Arabidopsis thaliana.  
PN WO200216655-A2.  
XX 28-FEB-2002.  
PD 24-AUG-2001; 2001WO-US026685.  
PF 24-AUG-2000; 2000US-0227866P.  
XX 24-AUG-2000; 2000US-0227866P.  
PR 26-JAN-2001; 2001US-0264647P.  
PR 22-JUN-2001; 2001US-0300111P.  
XX (SCRI) SCRIPPS RES INST.  
PA (SYGN) SYNGENTA PARTICIPATIONS AG.  
XX

PI Harper JF, Kreps J, Wang X, Zhu T;  
XX WPI; 2002-304127/34.  
DR Identifying a stress condition to which a plant cell has been exposed and  
XX producing plants with increased tolerance to these abiotic stresses.  
PT Claim 144; SEQ ID NO 893; 577bp + Sequence Listing; English.  
XX The invention relates to identifying a stress condition to which a plant  
CC cell has been exposed, comprising: (a) contacting nucleic acid  
CC representative of expressed polynucleotides in the plant cell with an  
CC array or probes representative of the plant cell genome; and (b)  
CC detecting a profile of expressed polynucleotides in the plant cell  
CC characteristic of a stress response. The method is useful in the  
CC production of transgenic plants, cells and seeds and in producing plants  
CC with increased tolerance to abiotic stress. The present sequence is that  
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
CC in methods of the invention. Note: The sequence data for this patent is  
CC not represented in the printed specification but is based on sequence  
CC information supplied to Derwent by the European Patent Office  
XX  
SQ Sequence 966 BP; 280 A; 157 C; 261 G; 268 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 1.02e-126 Length: 966  
Score: 1374.00 Matches: 267  
Percent Similarity: 91.08% Conservative: 29  
Best Local Similarity: 82.15% Mismatches: 25  
Query Match: 84.66% Indels: 4  
DB: 6 Gaps: 1  
US-09-931-457A-31 (1-325) x ABZ13088 (1-966)  
QY 1 MetalValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrPro 20  
DB 1 ATGGCTTCAGTGGCTCCTAAATTCGCTAAAGATGTAAGTAACTAATAGGAGACACTCCA 60  
QY 21 LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40  
DB 61 TTAGGTATCTAAACAAAGTAGCTTAAGATTGTGTTGTCATGTTGCTGCTAAGCTTGA 120  
QY 41 LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60  
DB 121 ATGATGGAACCATGTTCTAGCGTCNAAGACAGAGATTGTTATATGATGATGCTGATGCA 180  
QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80  
DB 181 GAAGCTAAAGAGACTTATTAAACCGGAGAGAGAGTGTGATGAGCCCAACAGTGGGAAC 240  
QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMet 100  
DB 241 ACTGAGTGGTTTACATTAATCTGACGCTGCAAGGCGTATTAAGCTTGTATTAACAATG 300  
QY 101 ProLaserMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120  
DB 301 CCAGCTCAATGATATAGAGAGAAATCATCTCTTAGCTTTTGAGCTGAGCTGAGTTGATT 360  
QY 121 LeuThrAspProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluGlnIleLeuAla 140  
DB 361 CTACTGATCCGGCTAAAGGTGAAAGGTGCTGTTGCAAGCGGAGAGATTTGGCG 420  
QY 141 LysThrProAsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHis 160  
DB 421 AAAACACCTAATGTTATATAGCTCAACAGTTGAGAAATCTGCTAATCCAAAGATCCAT 480  
QY 161 TyrGluThrThrGlyProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheVal 180  
DB 481 TATGAGACTACTGACCTGAGATATGAAAGGTTCTGTTGGAAGAAAGTGATGCTTTGTT 540  
QY 181 SerGlyIleGlyThrGlyLysThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsn 200  
DB 541 TCTGTATGTGTAAGT-----GGTGTGGGAGATATCTCAAGGAACAGAAC 588

QY 201 ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSergIlyLys 220  
DB 589 ACAACATMAAGCTGTATGCTGAGCCTGTGTAAGCCCTATTCTGTCGGTGGAAAG 648  
QY 221 ProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240  
DB 649 CCAGGTCCCAAAATTCAGGTATAGGTGCTGTTTATTCAGGCATTGGAATGTT 708  
QY 241 AsnLeuLeuAspGluValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260  
DB 709 GATCTTATAGATGAAGTTGTTCAAGTTTCAAGTGAAGAATCCATTGACATGGCAAGCTT 768  
QY 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSergIlyAlaAlaAlaAlaAla 280  
DB 769 CTGCTAGGAGAGAGGCTGTAGTGGAAATCTCATCTGAGCAGCAGCTACTGACACA 828  
QY 281 PheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSer 300  
DB 829 ATCAAACTTGCAAGAGGCCAGAAATGCTGGAGCTTATAGTGGCGGTGTTCCGAGT 888  
QY 301 PheGlyGluArgTyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSer 320  
DB 889 TTGGGGGAAGATATCTATCGACTGTAAGTCTGTTGATGACAGCAAGAAAGAGCAGAGACT 948  
QY 321 MetThrPheGluPro 325  
DB 949 ATGACCTTTGAGCCC 963

## RESULT 6

ADA68423  
ID ADA68423 standard; DNA; 966 BP.

XX AC ADA68423;

XX DT 20-NOV-2003 (first entry)

XX DE Arabidopsis thaliana gene, SEQ ID 683.

XX KM Plant; bacterial infection; fungal infection; viral infection; ds.

XX OS Arabidopsis thaliana.

XX PN WO2003000898-A1.

XX PD 03-JAN-2003.

XX PF 22-JUN-2001; 2001WO-IB001105.

XX PR 22-JUN-2001; 2001WO-IB001105.

XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,  
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX DR WPI; 2003-175290/17.

PT Identifying at least one gene involved in plant resistance or response to  
PT pathogenic infection for conferring resistance or tolerance to a plant to  
PT bacterial, fungal or viral infection by determining or detecting plant  
PT gene expression.

XX PS Claim 6; SEQ ID NO 683; 899pp; English.

XX CC The present invention relates to a method (M1) for identifying genes  
XX involved in plant resistance or response to pathogenic infection. M1  
XX comprises identifying a gene whose expression is significantly altered in  
XX the incompatible interaction of plant gene expression relative to  
XX expression of the gene in an uninfected plant, in a mutant plant that  
XX does not express a gene associated with response to pathogenic infection,  
XX or in a corresponding incompatible or compatible interaction. (M1) is  
XX useful for conferring resistance to resistance or tolerance to a plant to

CC bacterial, fungal or viral infection. The present sequence was used to  
CC illustrate the invention.

XX SQ Sequence 966 BP; 280 A; 157 C; 261 G; 268 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	1,02e-126	Length:	966
Score:	1374.00	Matches:	267
Percent Similarity:	91.08%	Conservative:	29
Best Local Similarity:	82.15%	Mismatches:	25
Query Match:	84.66%	Indels:	4
DB:	7	Gaps:	1

US-09-931-457A-31 (1-325) x ADA68423 (1-966)

QY 1 MetAlaValGluArgSergIlyIleAlaLysAspValThrGluLeuIleGlyLysThrPro 20  
DB 1 ATGGCTTCAGTGGCTCCTAAATCGCTAAAGATGTAACTGAATTAATAGGGAACACTCCA 60  
QY 21 LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40  
DB 61 TTAGTGTATCTAAACAAGTAGCTAAAGATTGTGTGTCATGTCGTGCTAAAGCTTGAG 120  
QY 41 LeuMetGluProCysSerSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60  
DB 121 ATGATGGAACCATGTTCTAGCGTCAAGACAGATTGTTATATGATGATGCTGATGCA 180  
QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSergIlyAsn 80  
DB 181 GAAGCTAAAGAGACTTATTAACCGGAGAGAGTGTGATGAGCAACAAGTGGGAAC 240  
QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMet 100  
DB 241 ACTGAGTGGTTTAGCATTTACTGCGCTCAAGGCTATAAGCTTGTATTAACAATG 300  
QY 101 ProLaserMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120  
DB 301 CCAGCTTCATGAGTATAGAGAGAGATCATCTCTTAGCTTTGAGCTGAGTTGATT 360  
QY 121 LeuThrAspProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluGluIleLeuAla 140  
DB 361 CTACTGATCCGGCTTAAGGTATGTAAGCTGCTGTCGAAAGCGGAGAGAGATTGGCG 420  
QY 141 LysThrProAsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHis 160  
DB 421 AAAACACCTTAATGTTATATGCTTCAACAGTTTGAGAAATCCTGTAATCCAAGATCCAT 480  
QY 161 TyrGluThrThrGlyProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheVal 180  
DB 481 TATGAGACTACTGCACTGAGATATGAAAGCTTCTGCTGGAAGAGTGATGGCTTGT 540  
QY 181 SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsn 200  
DB 541 TCTGATATGTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 588  
QY 201 ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSergIlyLys 220  
DB 589 ACAACATMAAGCTGTATGCTGAGCCTGTGTAAGCCCTATTCTGTCGGTGGAAAG 648  
QY 221 ProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240  
DB 649 CCAGGTCCCAAAATTCAGGTATAGGTGCTGTTTATTCAGGCATTGGAATGTT 708  
QY 241 AsnLeuLeuAspGluValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260  
DB 709 GATCTTATAGATGAAGTTGTTCAAGTTTCAAGTGAAGAATCCATTGACATGGCAAGCTT 768  
QY 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSergIlyAlaAlaAlaAlaAla 280  
DB 769 CTGCTAGGAGAGAGGCTGTATGAGTGAATCTCATCTGAGCAGCAGCTACTGACACA 828  
QY 281 PheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSer 300

Db 829 ATCAACTTGCAGAGAGCCAGAAATGCTGGAAGCTTATAGTGGCGGTGTTCCGAGT 888  
Qy 301 PheGlyGluArgTyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSer 320  
Db 889 TTCGGGGAAGATATCTATCGACTGACTGTTGATGACAGAGAAAGAGGACAGACT 948  
Qy 321 MetThrPheGluPro 325  
Db 949 ATGACCTTTGAGCCC 963  
RESULT 7  
ADA69412 standard; DNA; 993 BP.  
XX  
AC ADA69412;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Rice gene, SEQ ID 2735.  
XX  
KM Plant; bacterial infection; fungal infection; viral infection; rice;  
KM gene; ds.  
XX  
OS Oryza sativa.  
XX  
PN WO2003000898-A1.  
XX  
PD 03-JAN-2003.  
XX  
PF 22-JUN-2001; 2001WO-IB001105.  
XX  
PR 22-JUN-2001; 2001WO-IB001105.  
XX  
PA (SYGN) SYNGENTA PARTICIPATIONS AG.  
XX  
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX  
DR WPI; 2003-175290/17.  
XX  
PT Identifying at least one gene involved in plant resistance or response to  
PT pathogenic infection for conferring resistance or tolerance to a plant to  
PT bacterial, fungal or viral infection by determining or detecting plant  
PT gene expression.  
XX  
PS Claim 6; SEQ ID NO 2735; 899bp; English.  
XX  
CC The present invention relates to a method (M1) for identifying genes  
CC involved in plant resistance or response to pathogenic infection. M1  
CC comprises identifying a gene whose expression is significantly altered in  
CC the incompatible interaction of plant gene expression relative to  
CC expression of the gene in an uninfected plant, in a mutant plant that  
CC does not express a gene associated with response to pathogenic infection,  
CC or in a corresponding incompatible or compatible interaction. (M1) is  
CC useful for conferring resistance to resistance or tolerance to a plant to  
CC bacterial, fungal or viral infection. The present sequence was used to  
CC illustrate the invention.  
XX  
SQ Sequence 993 BP; 263 A; 207 C; 290 G; 233 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 2.6e-125 Length: 993  
Score: 1360.00 Matches: 260  
Percent Similarity: 92.11% Conservative: 32  
Best Local Similarity: 82.02% Mismatches: 25  
Query Match: 83.80% Indels: 0  
DB: 7 Gaps: 0  
US-09-931-457a-31 (1-325) x ADA69412 (1-993)  
Qy 8 IleAlaLysAspValThrGluLeuIleGlyLysThrProLeuValTyrLeuAsnLysLeu 27  
Db 40 ATGCCCAAGAGATGTACCGAGTGTATTGGGAAACAGCGCGTTGTGTACTCAACCGGTTG 99

Qy 28 AlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluProCysSerSer 47  
Db 100 ACCGATGGGTGCGTGGGCGCGTGGGCGGCAAGCTCGAGCTCCATGAGCCATGCTCCAGC 159  
Qy 48 ValLysAspArgIleGlyLysSerMetIleAlaAspAlaGluGluLysGlyLeuIleThr 67  
Db 160 GTCAAGCATAGATTGATGATCACTATGATCACTGATGACAGAGAGAGGCGGTGATCACT 219  
Qy 68 ProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIleGlyLeuAlaPhe 87  
Db 220 CCAAGCAAGAGTGTGCTGATGACCACTAGTGGCAACACAGCATTTGACTGGCCTTC 279  
Qy 88 MetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProAlaSerMetSerLeuGlu 107  
Db 280 ATGGCTGCTCAAAAGGCTTACAGGCTTGTACTCAGATGCCGCTCCATGAGCATAGAG 339  
Qy 108 ArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspProAlaLysGly 127  
Db 340 AGGAGATCATATTGAAGGCTTTTGTGCTGATTTGATTAATTAATTAATTAATTAATTA 399  
Qy 128 MetLysGlyAlaValGluLysAlaGluGluIleLeuAlaLysThrProAsnAlaTyrIle 147  
Db 400 ATGAAGAGAGCTGTCCAAAGGAGAGAAAGACTGGACAGCAAGCAAAACATCATTTATC 459  
Qy 148 LeuGluGluPheGluAsnProAlaAsnProLysValHisTyrGluThrThrGlyProGlu 167  
Db 460 CTCACAAATTCGAGAACCTTGCTTAACCAAGATCCATTACAGAGACCACTGAGCTGAA 519  
Qy 168 IleThrLysGlySerAspGlyLysIleAspAlaPheValSerGlyIleGlyThrGlyGly 187  
Db 520 ATCTGAAAGAGAACAGAGGTAAAGTTAGTGTATTGTTCTGCTATTGGACAGGTGC 579  
Qy 188 ThrIleThrGlyAlaGlyLysThrLeuLysGluGluAsnProAsnIleLysLeuIleGly 207  
Db 580 ACTATTACTGAGAGCTGACGATCACTCAGAGAGCAAAATCTGATATCAAGATCTATGCT 639  
Qy 208 ValGluProValGluSerProValLeuSerGlyLysProGlyProHisLysIleGlu 227  
Db 640 GTGAGCCAGTCGAGAGCGCTGCTTATCTGTGAAAGCTGGGCCACACAGATTCA 699  
Qy 228 GlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeuLeuAspGluValVal 247  
Db 700 GGAATTGAGAGCTGCTTTGTTCTCTGGGGTCTCGATGTGAACCTCATCAATGAACCTGA 759  
Qy 248 GlnIleSerSerAspGluAlaIleGluThrAlaLysLeuLeuAlaLeuLysGluGlyLeu 267  
Db 760 CAAGTTCAAGTATGATGAAGCTATCAAGATGCAAGAGGCTCTTGATGAAGAAGGTTG 819  
Qy 268 PheValGlyIleSerSerGlyAlaAlaAlaAlaAlaAlaPheGlnIleAlaLysArgPro 287  
Db 820 CTGGTTGAAATATCTTCAAGTGCAGCTGCAAGCAGCAGCTGTTAGGCTCGCTCAGAGCGCG 879  
Qy 288 GluAsnAlaGlyLysLeuIleValAlaValPheProSerPheGlyGluArgTyrLeuSer 307  
Db 880 GAAATGAGAGAAACTTTTGTGTGTCTTCCCAAGCTTTGTTGATGAGCGGTACCTTCG 939  
Qy 308 SerValLeuPheGluSerValArgArgGluAlaGluSerMetThrPheGlu 324  
Db 940 TCGGTGCTCTTCCAGTCCATCAAGAGAAAGCTGAAACATGTGTGTGA 990  
RESULT 8  
ADA70437 standard; DNA; 5259 BP.  
XX  
AC ADA70437;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Rice gene, SEQ ID 3760.  
XX  
KM Plant; bacterial infection; fungal infection; viral infection; rice;  
KM gene; ds.



XX Oryza sativa.  
OS WO2003000898-A1.  
XX PD 03-JUN-2003.  
XX PF 22-JUN-2001; 2001WO-IB001105.  
XX PR 22-JUN-2001; 2001WO-IB001105.  
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
PI Katagiri F, Qian S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX WPI; 2003-175290/17.  
XX Identifying at least one gene involved in plant resistance or response to  
PT pathogenic infection for conferring resistance or tolerance to a plant to  
PT bacterial, fungal or viral infection by determining or detecting plant  
PT gene expression.  
XX  
PS Claim 6; SEQ ID NO 3760; 899bp; English.  
XX  
CC The present invention relates to a method (M1) for identifying genes  
CC involved in plant resistance or response to pathogenic infection. M1  
CC comprises identifying a gene whose expression is significantly altered in  
CC the incompatible interaction of plant gene expression relative to  
CC expression of the gene in an uninfected plant, in a mutant plant that  
CC does not express a gene associated with response to pathogenic infection,  
CC or in a corresponding incompatible or compatible interaction. (M1) is  
CC useful for conferring resistance to resistance or tolerance to a plant to  
CC bacterial, fungal or viral infection. The present sequence was used to  
CC illustrate the invention.  
XX  
SQ Sequence 5259 BP; 1631 A; 1084 C; 1282 G; 1261 T; 0 U; 1 Other;  
  
Alignment Scores:  
Pred. NO.: 2.96e-124 Length: 5259  
Score: 1359.00 Matches: 260  
Percent Similarity: 89.72% Conservative: 28  
Best Local Similarity: 81.00% Mismatches: 33  
Query Match: 83.73% Indels: 0  
DB: 7 Gaps: 0  
  
US-09-931-457A-31 (1-325) x ADA70437 (1-5259)  
  
QY 1 MetAlaValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrPro 20  
DB 1 ATGGCGAATCAGGCCAATCGATTGCGAGTGTGACTGCACTGATCGCAATACTCC 60  
QY 21 LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40  
DB 61 CTGGTGTACCTGAACAGGTGTGTGATGTGAGGCCCAATCGCTGCCAAGCTTGAG 120  
QY 41 LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60  
DB 121 ATCATGAGCCCTGCTCACTGTCAGAGTATGATGTATAGCATGATTAAGATGCA 180  
QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80  
DB 181 GAAAGAGAAGGCTCTCACTACTCCAGAAAGAGTGTATTGATTGAGCTACTAGTGAAT 240  
QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMet 100  
DB 241 ACTGGTATTGGCTTGTCTTATGAGTGTGCTGCCAAGGTTATAGCTTAATCTGCAATG 300  
QY 101 ProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120  
DB 301 CCAGCATCATGAGCATGAGAAAGAGGATCATCTTAAGGCTTCGGTGCAGGTGTC 360  
QY 121 LeuThrAspProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluGlnIleLeuAla 140

DB 361 CTTACCGATCCTTCTGGGATGAAGGTGCAATCCAGAAAGGACAGTACCTGGCAGCA 420  
QY 141 LysThrProAsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHis 160  
DB 421 AAGATGCTTAATCTTACATTCTCCAGAGTTGAGAACCTTGCAACCCAAAGATCCAC 480  
QY 161 TyrGluThrThrGlyProGluIleThrLysGlySerAspGlyLysIleAspAlaPheVal 180  
DB 481 TATGAGACAACTGAGCTGAGATTGAGAGCCACTGCTGGAAAGTTGATATCCTTGTG 540  
QY 181 SerGlyIleGlyThrGlyGlyTyrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsn 200  
DB 541 TCTGCAATGGAACCGGTGTGACAGTAACCGAAGCAAGAAAGTACTCAAGAGCAGAAC 600  
QY 201 ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLys 220  
DB 601 CTTGAGATCAAGATCTATGTGTGTAACCGACAGAAAGTGAATTTATCTGAGAGAGA 660  
QY 221 ProGlyProHisLysIleGlnGlyTyrIleGlyAlaGlyPheIleProGlyValLeuGluVal 240  
DB 661 CTGTGTCACCAAAATCCAAAGTATAGTGTGTTTGTCTCTGTTGTTACTGATGTC 720  
QY 241 AsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260  
DB 721 AATCTCTGATGAAGTGTGTCCAGATCTCAAGTATGATGAGCCATGACATGCGAAGCAG 780  
QY 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla 280  
DB 781 CTAGCGTTGAAGAAGAGGCTGCTGTTGAGATCTCATCTGTTGCTGCTGCTGCTGCC 840  
QY 281 PheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSer 300  
DB 841 ATCAGGCTGCGCAGCGCCCGAGAACAAAGGAAGCTCGTAGTGTGTCTTCCCGAGC 900  
QY 301 PheGlyGluArgTyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSer 320  
DB 901 TTGGTGAAGCTACTGCTGTCTGTCTCTTCTGAAATGATCAAGAGGAGGCCGAAAC 960  
QY 321 Met 321  
DB 961 ATG 963  
  
RESULT 9  
AAQ74413 standard; cDNA to mRNA; 1483 BP.  
ID AAQ74413  
XX AC AAQ74413;  
XX 12-JUN-1995 (first entry)  
XX DE Spinach cysteine synthase.  
XX KW Cysteine; spinach; spinacia oleracea; pUC19; M13mp18; plant; nutrition;  
XX feed value; pharmaceutical; food additive; cosmetic; E.coli; ds.  
XX OS Spinacia oleracea.  
XX  
XX Key Location/Qualifiers  
XX CDS 62..1213  
XX FT /\*tag= a  
XX FT /product= "cysteine synthetase"  
XX PN JP06245773-A.  
XX PD 06-SEP-1994.  
XX XX  
XX 26-FEB-1993; 93JP-00038527.  
XX XX  
XX 26-FEB-1993; 93JP-00038527.  
XX XX  
XX (MITS ) MITSUBISHI CORP.  
XX (MITU ) MITSUBISHI KASEI CORP.  
XX PA

XX WPI, 1994-321282/40.  
DR P-PSDB; AAR63756.  
XX  
PT Gene encoding a cysteine synthesis enzyme - useful for elevating cysteine  
PT content in plant bodies increasing their nutritional value.  
XX  
PS Claim 2; Page 4-6; 6pp; Japanese.

XX The nucleotide sequence of the novel cysteine synthase enzyme from  
CC spinach, *Spinacia oleracea*. The probes (AAQ74414-5) were used to obtain a  
CC clone from a cDNA library derived from spinach seedling leaves. The 1.5  
CC kb insert was ligated into the cloning vectors pUC19 and M13mp18. The  
CC gene, 1483 bp, encodes a protein of 383 a.a. The cysteine gene is  
CC expressed in plants to elevate the cysteine content in the plant. The  
CC nutritional and feed value of the plant are expected to be enriched.  
CC Cysteine, one of the S-containing amino acids, can be used as materials  
CC for various pharmaceuticals, food additives or cosmetics. Production of  
CC the protein is a useful step in the production of cysteine

XX Sequence 1483 BP; 440 A; 272 C; 354 G; 417 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.97e-116 Length: 1483  
Score: 1271.00 Matches: 234  
Percent Similarity: 89.66% Conservative: 52  
Best Local Similarity: 73.35% Mismatches: 33  
Query Match: 78.31% Indels: 0  
DB: 2 Gaps: 0

US-09-931-457A-31 (1-325) x AAQ74413 (1-1483)

QY 3 ValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrProLeuVal 22  
DB 245 ATGAAGGCGCTTAACATTGCTGAAGATGTTCTCAGCTAATTGAAAAAACCCTGATG 304  
QY 23 TyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGluMet 42  
DB 305 TATCTCAACAATGATCAAAAGATCAGTTGCAACATTCGCGAAGCTTGAGACATG 364  
QY 43 GluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGlu 62  
DB 365 GAACCTTGCTGAGTGTCAAGACAGATGCTGCTACATGATGATGATGATGATGATG 424  
QY 63 LysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGly 82  
DB 425 AAAGAGATTATCAACCTGGAAGAACTACTCTAGTGAGCCTTACGAGTGGATACTGCA 484  
QY 83 IleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProAla 102  
DB 485 ATAGACTTGCTTATAGTCTGCTGCCAGAGATACAGATTAACCTTGACATGCCAGCT 544  
QY 103 SerMetSerLeuGluArgGlyIleLeuLeuAlaPheGlyAlaGluLeuValLeuThr 122  
DB 545 TCCATGAGTATGGAAGGAGATGATCTTGAAGCATTTGAGAGCTGAGTGTGCTGACT 604  
QY 123 AspProAlaLysGlyMetLysGlyAlaValGluLysAlaGluGluIleLeuAlaLysThr 142  
DB 605 GATCCAGCTAAGGGAATGAAGAGACAGTGAAGAGCTGAAGAAATTTGAGAAACT 664  
QY 143 ProAsnAlaTyrIleLeuGlnIlePheGluAsnProAlaAsnProLysValHisTyrGlu 162  
DB 665 CCGATTCTCTACATGCTTCAAGCTTGAACAATCTGCAATCCCAAGATACATTAAGAG 724  
QY 163 ThrThrGlyProGluIleThrLysGlySerAspGlyLysIleAspAlaPheValSerGly 182  
DB 725 ACAACAGGTCCCGAGATCTGGGAAGACACAAAAGGCAAGTGAACATTTTGTCCAGGC 784  
QY 183 IleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGluAsnProAsn 202  
DB 785 ATTGGAAGTGGAGAACGATTTCTGAGTGGACGCTACCTCAAGAACGTAACCTCTGCT 844  
QY 203 IleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLysProGly 222

DB 845 GTCCAGTAATTGGTATAGAACCTACAGAAAGCAACATACCTTCTGTGGAAGCCTGCT 904  
QY 223 ProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeu 242  
DB 905 CCACACAGATTCAAGGACTTGAGCTGCTGCTTTGTTCCAGCAATTAGATTGGGTGTG 964  
QY 243 LeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeuAla 262  
DB 965 ATGATGAAGTTATAGAGGATATCTAGTGAAGAGCTGTAGAAATGCCAAATTGGCA 1024  
QY 263 LeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaPheGln 282  
DB 1025 ATGAAGAGGCTTGTGTGTTGGCATTTCTGAGCAGCAGCAGCTGCTGCACTCAGG 1084  
QY 283 IleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSerPheGly 302  
DB 1085 ATGGTAAGAAGCTGAAATGCAAGAAACTTATTGCTGTTGTTCCCAAGCTTTGCT 1144  
QY 303 GluArgTyrLeuSerSerValLeuPheGluSerValArgGluAlaGluSerMet 321  
DB 1145 GAGAGATACCTGTCATTCATTGTTGCTCAGCTATTGAGAGAGGTGTGAACAATG 1201

RESULT 10

AAC34662 standard; DNA; 1503 BP.

AC AAC34662;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 7429.

XX Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway; metabolic pathway;

KW promoter; termination sequence; ss.

XX Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.

Mon May 24 08:18:51 2004

us-09-931-457a-31.p2n.rng

Page 11

PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 20-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 21-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 10-AUG-1999; 99US-0147935P.  
PR 11-AUG-1999; 99US-0148171P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155559P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158369P.  
PR 12-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 22-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.

PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161320P.  
PR 28-OCT-1999; 99US-0161392P.  
PR 28-OCT-1999; 99US-0161392P.  
PR 28-OCT-1999; 99US-0161393P.  
PR 29-OCT-1999; 99US-0162142P.

## Alignment Scores:

Pred. No.:	2.97e-115	Length:	1503
Score:	1261.00	Matches:	239
Percent Similarity:	87.77%	Conservative:	41
Best Local Similarity:	74.92%	Mismatches:	39
Query Match:	77.70%	Indels:	0
DB:	3	Gaps:	0

US-09-931-457a-31 (1-325) x AAC34662 (1-1503)

QY 3 ValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrProLeuVal 22  
DB 363 GTTGAAGGGCTCAATATCCCGGATAACGCCGCTGAGCTTATTTGGAAACTCCGATGGTG 422  
QY 23 TyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuLeuMet 42  
DB 423 TACTTGACAAATGATGATCAAGGGCTGTGTGCAAGTGTGCTGCTAAGCTTGATCATG 482  
QY 43 GluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGlu 62  
DB 483 GAACCATGTTGCAAGTGTCAAGATAGATTGGGTACAGTATGATTACTGATGCTGAAGAG 542  
QY 63 LysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGly 82  
DB 543 AAAGGACTTATTAACACCTGGAAGAAGTGTCTTGTGGAATCTACGAGTGGGAACACAGGG 602  
QY 83 IleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIlePheMetProAla 102  
DB 603 ATTGGCCTTGCAATCAATGCTGCTCAAAAGGCTATTAACCTTATCTTGACCATGCTGCG 662  
QY 103 SerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThr 122  
DB 663 TCCATGAGTTTGAAGGGGGGTTCTTTGAGGGCATTTGGAGCTGAGCTTGTTGTTAACT 722  
QY 123 AspProAlaLysGlyMetLysGlyAlaValAlaGlnLysAlaGluGluIleLeuAlaLysThr 142  
DB 723 GAACCTGCAAAAGGTATGACTGAGCAATTCAGAGCTGAAGAAATCTTGAAAAAAACT 782  
QY 143 ProAsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGlu 162  
DB 783 CCCAATTCCTACATGCTTCAACAGTTTGACCAACCTGCAATCCCAAGCTTATATGAG 842  
QY 163 ThrThrGlyProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheValSerGly 182  
DB 843 ACGACTGCTCTGAGATTGGGAAGATACACAAGAGCAAAATCGATATATTTGGCGGG 902  
QY 183 IleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsnProAsn 202  
DB 903 ATTGAAGCTGTGGAAGCTATCACTGTGTGTCATTCATTAAAGAAAGAAAACTGAA 962  
QY 203 IleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLysProGly 222  
DB 963 TTGAAGTTATTTGTCGAACCCACGAAAGTCTATACTTTCTGTGGAACCCGGA 1022  
QY 223 ProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeu 242  
DB 1023 CCTCACAAGATTCAGAGATTTGAGCTGATTTGTAAACCAAGATTGATCTGCTATT 1082  
QY 243 LeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeuLeuAla 262  
DB 1083 GTAGATGATATACATAGCATTTCCAGTGAAGAGCTATTGAACCTCGAAGCACTAGCT 1142  
QY 263 LeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAlaPheGln 282  
DB 1143 CTCGAGGAGGCTTGTGTTGTTGTTATATCTTCTGAGCTGCTGCTGCTGCAATCCAG 1202

QY 283 IleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSerPheGly 302  
DB 1203 GTTGAAGAGAGACTGAATAATGCCGGAACCTCATAGCCGTTGTCTCCGAGCTTCGGG 1262  
QY 303 GluArgTyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSerMet 321  
DB 1263 GACGTTACTCTCGACCCAGCTTTTCCAGTCGATTGAGAGAGTGCAGCAATG 1319

## RESULT 11

ABZ12294  
ID ABZ12294 standard; DNA; 1329 BP.

AC ABZ12294;

DT 21-JAN-2003 (first entry)

DE Arabidopsis thaliana stress regulated gene SEQ ID NO 99.

KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

OS Arabidopsis thaliana.

PN W0200216655-A2.

PD 28-FEB-2002.

PF 24-AUG-2001; 2001WO-US026685.

PR 24-AUG-2000; 2000US-0227866P.

PR 26-JAN-2001; 2001US-0264647P.

PR 22-JUN-2001; 2001US-0300111P.

XX (SCRI ) SCRIPPS RES INST.

XX (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX Harper JF, Kreps J, Wang X, Zhu T;

XX WPI; 2002-304127/34.

PT Identifying a stress condition to which a plant cell has been exposed and

PT producing plants with increased tolerance to these abiotic stresses.

XX Claim 144; SEQ ID NO 99; 577bp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant

XX cell has been exposed, comprising: (a) contacting nucleic acid

XX representative of expressed polynucleotides in the plant cell with an

XX array or probes representative of the plant cell genome; and (b)

XX detecting a profile of expressed polynucleotides in the plant cell

XX characteristic of a stress response. The method is useful in the

XX production of transgenic plants, cells and seeds and in producing plants

XX with increased tolerance to abiotic stress. The present sequence is that

XX of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used

XX in methods of the invention. Note: The sequence data for this patent is

XX not represented in the printed specification but is based on sequence

XX information supplied to Derwent by the European Patent Office

XX SQ Sequence 1329 BP; 361 A; 268 C; 333 G; 367 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	9.92e-115	Length:	1329
Score:	1255.00	Matches:	241
Percent Similarity:	84.02%	Conservative:	43
Best Local Similarity:	71.30%	Mismatches:	36
Query Match:	77.33%	Indels:	18
DB:	6	Gaps:	2

US-09-931-457a-31 (1-325) x ABZ12294 (1-1329)

QY 2 AlaValGluArgSer-----GlyIleAlaLysAspValThrGluLeu 15  
DB 301 GCTGTTAAGCGGAGAGACTGAGCTGATGACTCAACATGCTGACAAATGTCTCAGCTC 360



QY 16 IleglylsThrProleuValIleuAsnLysLeuAlaaspGlyCysValAlaargVal 35  
 Db 361 ATGGGAAACTCTCTATGCTGACTGACACAGTATAGCAAAAGTTGTGTAGCAAAATATT 420  
 QY 36 AlaAlaLysLeuGluLeuMetGluProCysSerSerValLysAspArgIleGlyTyrSer 55  
 Db 421 GCTGCCAACTCGAGATCATGAGCCATGTTGCAAGTCAAGAGATAGGATTGGTACAGT 480  
 QY 56 MetIleAlaAspAlaGluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGlu 75  
 Db 481 ATGGTTACTGATGCTGAGCAAAAGATTATATCACTGGAAGAGTGTCTTGTGAA 540  
 QY 76 ProThrSerGlyAsnThrGlyIleGlyLeuAlaPheMetAlaAlaArgGlyTyrLys 95  
 Db 541 CTTACAGTGAACACACGGGTATTCGCTAGCATTCATCGCTCTTCAAGGGCTATAGA 600  
 QY 96 LeuIleIleThrMetProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPhe 115  
 Db 601 CTTATCTTGACATGCGCAGATCATGAGTATGGAAGAAGGGTCTTTTAAAGGCAATT 660  
 QY 116 GlyAlaGluLeuValLeuThrAspProAlaLysGlyMetLysGlyAlaValGluLysAla 135  
 Db 661 GGAGCTGAGCTTGATTTAACAAGATCTGCTTAAAGGCATGAGCGGTTGAGAGGCT 720  
 QY 136 GluGluIleLeuAlaLysThrProAsnAlaTyrIleLeuGluGlnPheGluAsnProAla 155  
 Db 721 GAAGAGATCTTGAAAAACACTCTGATGATCATGCTCCACAGTTTGACAAACCTGCA 780  
 QY 156 AsnProLysValHisTyrGluThrThrGlyProGluIleTrpLysGlySerAspGlyLys 175  
 Db 781 AATCCAAAGATTCTATTACGAGACTAGTCTGCTGAGATCTGGATGATACAAAGCGCAAG 840  
 QY 176 IleAspAlaPheValSerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyr 195  
 Db 841 GTGATATCTTTGTGAGGAATTGGAAGTGTGGAATCATGCTGATGCTGATGCTGATTT 900  
 QY 196 LeuLysGluGlnAsnProAsnIleLysLeuIleGlyValGluProValGluSerProVal 215  
 Db 901 ATTAAGAGAAATAATCCCAAAACGAGGTTATGTGTAGAACCCACGAAAGCGATATA 960  
 QY 216 LeuSerGlyLysProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIlePro 235  
 Db 961 CTTCTGTGGCAACCGGAGCTCACAGATTCAGGAATTGGAAGCTGGAATTTATACCT 1020  
 QY 236 GlyValLeuGluValAsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIle 255  
 Db 1021 AAGAATTTGATCAGAAATTTATGATGAAGTCAAGTCAAGTGAAGAGCTATA 1080  
 QY 256 GluThrAlaLysLeuLeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAla 275  
 Db 1081 GAAACCCGCAAGCACTAGCTCTTAAGGAAGCTTGATGTTGTTATATCATCTGAGCT 1140  
 QY 276 AlaAlaAlaAlaAlaPheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleVal 295  
 Db 1141 GCACTGCTGCTGCAATCAAGTGGCTAAGAGACCTGAGAAATGCTGAAAACCTCATAGCT 1200  
 QY 296 -----AlaValPheProSerPheGlyGlu 303  
 Db 1201 GTAAGTTTATGATGATGATGATGCGGTTTTCAGGTTGTTTCCCGAGCTTGGGGAG 1260  
 QY 304 ArgTyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSerMet 321  
 Db 1261 AGATATCTATGACCCCGCTGTTCCAGTCTATCCGGGAAGAGTGGAGAAATATG 1314

RESULT 12  
 AAC40774  
 ID AAC40774 standard; DNA; 1224 BP.  
 AC AAC40774;  
 XX 17-OCT-2000 (first entry)  
 DT  
 XX Arabidopsis thaliana DNA fragment SEQ ID NO: 29501.

XX Hybridisation assay; genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway; metabolic pathway;  
 KW promoter; termination sequence; ss.  
 OS Arabidopsis thaliana.  
 PN EP1033405-A2.  
 XX 06-SEP-2000.  
 PD 25-FEB-2000; 2000EP-00301439.  
 PF 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 23-APR-1999; 99US-0130891P.  
 PR 28-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 04-MAY-1999; 99US-0132484P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 06-MAY-1999; 99US-0132487P.  
 PR 07-MAY-1999; 99US-0132863P.  
 PR 11-MAY-1999; 99US-0134256P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138847P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139492P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.

PR	23-JUN-1999;	99US-01403533P.
PR	23-JUN-1999;	99US-01403544P.
PR	24-JUN-1999;	99US-01406955P.
PR	28-JUN-1999;	99US-01408233P.
PR	29-JUN-1999;	99US-01409911P.
PR	30-JUN-1999;	99US-01412872P.
PR	01-JUL-1999;	99US-01418427P.
PR	01-JUL-1999;	99US-01421544P.
PR	02-JUL-1999;	99US-01420553P.
PR	06-JUL-1999;	99US-01423909P.
PR	08-JUL-1999;	99US-01428203P.
PR	12-JUL-1999;	99US-01429777P.
PR	13-JUL-1999;	99US-01435442P.
PR	14-JUL-1999;	99US-01436244P.
PR	15-JUL-1999;	99US-01440055P.
PR	16-JUL-1999;	99US-01440855P.
PR	16-JUL-1999;	99US-01440866P.
PR	19-JUL-1999;	99US-01443325P.
PR	19-JUL-1999;	99US-01443331P.
PR	19-JUL-1999;	99US-01443332P.
PR	19-JUL-1999;	99US-01443333P.
PR	19-JUL-1999;	99US-01443334P.
PR	19-JUL-1999;	99US-01443352P.
PR	19-JUL-1999;	99US-01446322P.
PR	20-JUL-1999;	99US-01448844P.
PR	20-JUL-1999;	99US-01448844P.
PR	21-JUL-1999;	99US-01448814P.
PR	21-JUL-1999;	99US-01450866P.
PR	22-JUL-1999;	99US-01450885P.
PR	22-JUL-1999;	99US-01450887P.
PR	22-JUL-1999;	99US-01450897P.
PR	23-JUL-1999;	99US-01451521P.
PR	23-JUL-1999;	99US-01451545P.
PR	23-JUL-1999;	99US-01452218P.
PR	23-JUL-1999;	99US-01452224P.
PR	26-JUL-1999;	99US-01455913P.
PR	27-JUL-1999;	99US-01455913P.
PR	27-JUL-1999;	99US-01455918P.
PR	28-JUL-1999;	99US-01459519P.
PR	02-AUG-1999;	99US-01463866P.
PR	02-AUG-1999;	99US-01463888P.
PR	03-AUG-1999;	99US-01470388P.
PR	04-AUG-1999;	99US-01472044P.
PR	04-AUG-1999;	99US-01473022P.
PR	05-AUG-1999;	99US-01477192P.
PR	06-AUG-1999;	99US-01473033P.
PR	06-AUG-1999;	99US-01474161P.
PR	09-AUG-1999;	99US-01474933P.
PR	09-AUG-1999;	99US-01479355P.
PR	10-AUG-1999;	99US-01481371P.
PR	11-AUG-1999;	99US-01483191P.
PR	12-AUG-1999;	99US-01483411P.
PR	13-AUG-1999;	99US-01486844P.
PR	16-AUG-1999;	99US-01493368P.
PR	17-AUG-1999;	99US-01491755P.
PR	18-AUG-1999;	99US-01497222P.
PR	20-AUG-1999;	99US-01497222P.
PR	20-AUG-1999;	99US-01497233P.
PR	20-AUG-1999;	99US-01499292P.
PR	23-AUG-1999;	99US-01499302P.
PR	23-AUG-1999;	99US-01505666P.
PR	26-AUG-1999;	99US-01508844P.
PR	27-AUG-1999;	99US-01510655P.
PR	27-AUG-1999;	99US-01510666P.
PR	30-AUG-1999;	99US-01513033P.

PR	31-AUG-1999;	99US-0151438P
PR	01-SEP-1999;	99US-0151930P
PR	07-SEP-1999;	99US-0152363P
PR	10-SEP-1999;	99US-0153070P
PR	13-SEP-1999;	99US-0153758P
PR	15-SEP-1999;	99US-0154018P
PR	16-SEP-1999;	99US-0154403P
PR	20-SEP-1999;	99US-0154779P
PR	22-SEP-1999;	99US-0155139P
PR	23-SEP-1999;	99US-0155496P
PR	24-SEP-1999;	99US-0155659P
PR	28-SEP-1999;	99US-0155658P
PR	29-SEP-1999;	99US-0156596P
PR	04-OCT-1999;	99US-0157117P
PR	05-OCT-1999;	99US-0157753P
PR	06-OCT-1999;	99US-0157865P
PR	07-OCT-1999;	99US-0158029P
PR	08-OCT-1999;	99US-0158432P
PR	12-OCT-1999;	99US-0158369P
PR	13-OCT-1999;	99US-0159293P
PR	13-OCT-1999;	99US-0159294P
PR	13-OCT-1999;	99US-0159295P
PR	14-OCT-1999;	99US-0159329P
PR	14-OCT-1999;	99US-0159330P
PR	14-OCT-1999;	99US-0159331P
PR	14-OCT-1999;	99US-0159637P
PR	14-OCT-1999;	99US-0159638P
PR	18-OCT-1999;	99US-0159584P
PR	21-OCT-1999;	99US-0160741P
PR	21-OCT-1999;	99US-0160767P
PR	21-OCT-1999;	99US-0160768P
PR	21-OCT-1999;	99US-0160814P
PR	21-OCT-1999;	99US-0160815P
PR	22-OCT-1999;	99US-0160980P
PR	22-OCT-1999;	99US-0160989P
PR	22-OCT-1999;	99US-0160989P
PR	25-OCT-1999;	99US-0161404P
PR	25-OCT-1999;	99US-0161405P
PR	25-OCT-1999;	99US-0161406P
PR	26-OCT-1999;	99US-0161359P
PR	26-OCT-1999;	99US-0161360P
PR	26-OCT-1999;	99US-0161361P
PR	28-OCT-1999;	99US-0161920P
PR	28-OCT-1999;	99US-0161922P
PR	28-OCT-1999;	99US-0161933P
PR	29-OCT-1999;	99US-0162142P

Alignment Scores:		
Pred. No.:	6.33e-109	length:
Score:	1196.00	Matches:
Percent Similarity:	84.74%	Conservative:
Best local Similarity:	71.34%	Mismatches:
Query Match:	73.69%	Indels:
DB:	3	Gaps:
		0

US-09-931-457A-31 (1-325) x AAC40774 (1-1224)

QY		4	GluArgSerGlyIleAlaIysAspValThrGluLeuIleGlyLysThrProLeuValTyr	23
Db		51	GATCGGTGCTTGATCAAGACGATGCACGTGAATTGATTGGTAACACCAACTGGTGTAT	110
QY		24	LeuAsnLysLeuAlaAspGlyCysValAlaIargValAlaAlaLysLeuGluLeuMetGlu	43
Db		111	CTGAACAATGTTGTTGATGGTGGCTGGCTCGTATCGCTGCAAGCTTCGAGATGATGAG	170
QY		44	ProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGluLys	63
Db		171	CCTTGTTCTAGCGCTCAAGAACAAGATCGCGTATAGTATGATCAAAGATGCAGAAAGAAA	230
QY		64	GlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIle	83
Db		231	GGAATTGATTACTCCCGAAGAGATACATTTGATAGAGCCAACCTGCTGGTAACACC	290

```
OY 84 GlyLeuAlaPheMetAlaAlaAlaArgGlyTyrTyrLeuIleIleThrMetProAlaSer 103
DB 291 GATTAGCTTGCATGGAGCTGCAAGAGGCTATAAAGTATCTTGATGCTTCAACT 350
OY 104 MetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAsp 123
DB 351 ATGAGCTTAGAGAGAGAAATCATCTGAGGGCACTAGCTGCAAGAGCTTCACTCTCGGAC 410
OY 124 ProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluGluIleLeuAlaLysThrPro 143
DB 411 CAGCGCATAGCCCTTAAGAGATGTGGAGAAAAGTGAAGCCGATTTTAAGCAAACTCCT 470
OY 144 AsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThr 163
DB 471 GGTGTTACATTCACACAAATTTGAAATCTCGAAACCCGAGATTCAATTACCGAAC 530
OY 164 ThrGlyProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheValSerGlyIle 183
DB 531 ACGGACCGGAAATATGAGAGATTCAGCCCGGAAAGTAGATATATGTGCTGCGCGTA 590
OY 184 GlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluAsnProAsnIle 203
DB 591 GGGACTGTGGAAGTCTGCTGAGTAGAGGAGTCTCAAGAGAGAGAGAGAGAGAGAGATC 650
OY 204 LysLeuIleGlyValGluProValGluSerProValLeuSerGlyLysProGlyPro 223
DB 651 AAGTTTGTGTGTGAGACCAAGTAGAAAGTCCGGTACTTAGCGAGGTCAACCGGTCCA 710
OY 224 HisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeu 243
DB 711 CATTTGATTCAAGGAAATGCTCTGTATCGTATCGTCCATTCAATTGAGCTTAACCATGTT 770
OY 244 AspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeuAlaLeu 263
DB 771 GATGAATATTATCAAGTGGCAGGTGAAGAGGCTATTGAACAGCCAGCTTCTGCCCTC 830
OY 264 LysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAlaPheGlnIle 283
DB 831 AAGAGAGATTACTGCTGGGAAATATCTCTGAGAGCCGAGAGAGAGAGAGAGAGT 890
OY 284 AlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSerPheGlu 303
DB 891 GCAAGCGCGCCAGAAACCGCGGAAACTCATTTGTGTGTTTCTTCTAGTGAAGAGAA 950
OY 304 ArgTyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSerMetThrPhe 323
DB 951 CGTTATTATCGACTAACTGTTGATTGATTGATATGAAAGCAGAGAAATTTGCCCTATT 1010
OY 324 Gln 324
DB 1011 GAA 1013
```

RESULT 13  
ABZ12690  
ID ABZ12690 standard; DNA; 972 BP.

XX AC ABZ12690;

XX DF 21-JAN-2003 (first entry)

XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 495.

XX KM Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX OS Arabidopsis thaliana.

XX PN WO20021655-A2.

XX PD 28-FEB-2002.

XX PF 24-AUG-2001; 2001WO-US026685.

XX

```
PR 24-AUG-2000; 2000US-0227866P.  
PR 26-JAN-2001; 2001US-0264647P.  
PR 22-JUN-2001; 2001US-0306111P.  
XX  
PA (SCRI ) SCRIPPS RES INST.  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX  
PI Harper JF, Kreps J, Wang X, Zhu T;  
XX  
DR WPI; 2002-304127/34.  
PT Identifying a stress condition to which a plant cell has been exposed and  
PT producing plants with increased tolerance to these abiotic stresses.  
XX  
PS Claim 144; SEQ ID NO 495; 577pp + Sequence Listing; English.  
XX  
CC The invention relates to identifying a stress condition to which a plant  
CC cell has been exposed, comprising: (a) contacting nucleic acid  
CC representative of expressed polynucleotides in the plant cell with an  
CC array or probes representative of the plant cell genome; and (b)  
CC detecting a profile of expressed polynucleotides in the plant cell  
CC characteristic of a stress response. The method is useful in the  
CC production of transgenic plants, cells and seeds and in producing plants  
CC with increased tolerance to abiotic stress. The present invention is that  
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ1574) used  
CC in methods of the invention. Note: The sequence data for this patent is  
CC not represented in the printed specification but is based on sequence  
CC information supplied to Derwent by the European Patent Office  
XX  
SQ Sequence 972 BP; 286 A; 178 C; 265 G; 243 T; 0 U; 0 Other;
```

#### Alignment Scores:

Pred. No.:	1,16e-108	Length:	972
Score:	1192.00	Matches:	227
Percent Similarity:	84.74%	Conservative:	45
Best Local Similarity:	70.72%	Mismatches:	49
Query Match:	73.44%	Indels:	0
DB:	6	Gaps:	0

US-09-931-457A-31 (1-325) x ABZ12690 (1-972)

```
OY 4 GluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrProLeuValTyr 23
DB 7 GATCGGTGCTGATCAAGAAGATATCACTGAATTGATTGTGAACACCAATGCTGAT 66
OY 24 LeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuLeuMetGlu 43
DB 67 CTGAACAATGTTGTTGATGTTGCTGCTGATCGCTGGAAGCTTGAAGATGATGAG 126
OY 64 GlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIle 83
DB 127 CCTGTTCTAGCTCAAGAAGATCGCTGATATGATGATCAAGATGACAGAACAA 186
OY 64 GlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIle 83
DB 187 GGATTGATTACTCCCGAAAGAGTACATGATAGAGCCAACTGCTGTAACACCGGATT 246
OY 84 GlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProAlaSer 103
DB 247 GGTTAGCTTGCATGGAGCTGCAAGAGCTATAAAGTATCTTGTATGCTTCAACT 306
OY 104 MetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAsp 123
DB 307 ATGAGCTTAGAGAGAGAAATCATCTGAGGGCACTAGTGACAGAGCTTCACTCTCGAC 366
OY 124 ProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluGluIleLeuAlaLysThrPro 143
DB 367 CAGCGCATAGGCTTAAAGAGATGTGGAGAAAAGTGAAGCGATTTTAAGCAAACTCCT 426
OY 144 AsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThr 163
DB 427 GGTGTTACATTCACACAAATTTGAAATCTCGAAACCCGAGATTCAATTACCGAAC 486
```

OY 164 ThrGlyProGluIleTrpLysGlySerAspGlyLysIleAspAlaPheValSerGlyIle 183  
DB 487 ACGGACCGGAATATGAGATTCAGCCGGGAAGATATATTGGTCCCTGGCGTA 546  
OY 184 GlYThrGlyGlyThrIleThrGlyAlaGlyLysIleLysGluGlnAsnProAsnIle 203  
DB 547 GCGACTGGTGAAGTCTACTGAGTAGGAGGAGTTCCTCAAGAGCAGACAAAGACATC 606  
OY 204 LysLeuIleGlyValGluProValGluSerProValLeuSerGlyLysProGlyPro 223  
DB 607 AAGTTTGTGTGGTGAACAGTAGAAGTCCGGTACTTAGCGGAGGTCAACCAAGTCCA 666  
OY 224 HisLysIleGlyGlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeu 243  
DB 667 CATTGATTCAAGGAATGGCTCTGGTATCGTCCCAATTGCAATTGCACTTAACCATTTT 726  
OY 244 AspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeuValAla 263  
DB 727 GATGAATTTATTCACGTGGCAGGTGAAGAGGCTATGAACAACAGCCAGCTTTGCCCTC 786  
OY 264 LysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAlaPheGlnIle 283  
DB 787 AAAGAGGATTAATCTGTGGGAATATCTCTGGAGCCGACAGCAGCGGCTGCGTTAAAGTT 846  
OY 284 AlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSerPheGlyGlu 303  
DB 847 GCAGAGCCGCCAGAAACGCCGCAACTCATTGTGTGCTTTTCTCTAGTGAAGAGAA 906  
OY 304 ArgTyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSerMetThrPhe 323  
DB 907 CGTATTATTCGACTAACTGTTCTGATTGATGATGAAGCAGAGAAATTGGCTATT 966  
OY 324 Glu 324  
DB 967 GAA 969  
RESULT 14  
AAC45719  
ID AAC45719 standard; DNA; 999 BP.  
AC AAC45719;  
AC 18-OCT-2000 (first entry)  
DT Arabidopsis thaliana DNA fragment SEQ ID NO: 47517.  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 47517.  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
OS Arabidopsis thaliana.  
XX EP1033405-A2.  
XX 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
XX 05-MAR-1999; 99US-0123180P.  
XX 09-MAR-1999; 99US-0123548P.  
XX 23-MAR-1999; 99US-0125788P.  
XX 25-MAR-1999; 99US-0126264P.  
XX 29-MAR-1999; 99US-0126785P.  
XX 01-APR-1999; 99US-0127462P.  
XX 06-APR-1999; 99US-0128234P.  
XX 08-APR-1999; 99US-0128714P.  
XX 16-APR-1999; 99US-0129845P.  
XX 19-APR-1999; 99US-0130077P.  
XX 21-APR-1999; 99US-0130449P.  
XX 23-APR-1999; 99US-0130510P.  
XX 23-APR-1999; 99US-0130891P.  
XX 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.





OY 224 HisIysIleGlnGlyIleGlyAlaGlyPheIleProGlyValIleuGluValAsnIleu 243  
 Db 694 CATTTGATTCAGGAAATGGCTCTGTATCGTCCCATTCATTGGACTTAACCATGTT 753  
 OY 244 AspGluValAlaGlnIleSerSerAspGluAlaIleGluThrAlaIleuAlaIleu 263  
 Db 754 GATGAATATTATCAAGTGGCAGGTGAAGAGGCTATTGAACAGCCAGCTTCCCTC 813  
 OY 264 LysGluGlyIleuPheValGlyIleSerSerGlyAlaAlaAlaAlaAlaPheGlnIle 283  
 Db 814 AAAGAAGATTACTGGTGGGAATATCTCTGGAGCCGACAGCGGCTGCTTAAGGTT 873  
 OY 284 AlAlaIleArgProGluAsnAlaGlyIleValAlaAlaValPheProSerPheGlyGlu 303  
 Db 874 GCAGAGCGGCGCAGAAACCGCGGAACTCATTTGTGCTTTTCTTGTGAGAGGAA 933  
 OY 304 ArgTyrIleuSerSerValIleuPheGluSerValArgGluAlaGluSerMetThrPhe 323  
 Db 934 CGTTATTATCGACTAACTGTTCATTGATTGATATGAGCAGAGAAATTGCTATT 993  
 OY 324 Glu 324  
 Db 994 GAA 996  
 RESULT 15  
 AAC41206  
 ID AAC41206 standard; DNA; 1423 BP.  
 AC AAC41206;  
 DT 17-OCT-2000 (first entry)  
 XX Arabidopsis thaliana DNA fragment SEQ ID NO: 31038.  
 DB Arabidopsis thaliana DNA fragment SEQ ID NO: 31038.  
 XX Hybridisation assay; genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway; metabolic pathway;  
 KM promoter; termination sequence; ss.  
 OS Arabidopsis thaliana.  
 XX Arabidopsis thaliana.  
 PN EP1033405-A2.  
 XX 06-SEP-2000.  
 PD 06-SEP-2000.  
 XX 25-FEB-2000; 2000EP-00301439.  
 PF 25-FEB-2000; 2000EP-00301439.  
 XX 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 23-APR-1999; 99US-0130891P.  
 PR 28-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 04-MAY-1999; 99US-0132407P.  
 PR 05-MAY-1999; 99US-0132484P.  
 PR 06-MAY-1999; 99US-0132485P.  
 PR 07-MAY-1999; 99US-0132486P.  
 PR 11-MAY-1999; 99US-0132833P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138847P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139492P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140991P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142920P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 20-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145088P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145913P.

PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.

PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

## Alignment Scores:

Pred. No.:	1.92e-108	length:	1423
Score:	1192.00	Matches:	227
Percent Similarity:	84.74%	Conservative:	45
Best Local Similarity:	70.72%	Mismatches:	49
Query Match:	73.44%	Indels:	0
DB:	3	Gaps:	0

US-09-931-457A-31 (1-325) x AAC41206 (1-1423)

QY 4 GluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrProLeuValTyr 23  
DB 246 GATCGGTGCTGATCAAGAACGATGTCACTGAATTGGTTGGTAACACCAATGTGTAT 305

QY 24 LeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGlu 43  
DB 306 CTGAACAATGTTGTGATGTTGCGTGGCTGCTATCGCTGCAAGCTTGAGATGATGAG 365

QY 44 ProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGluLys 63  
DB 366 CCTGTTCTAGCGCTCAAGACAGAAATCGCTATGATGATCAAGATGCAAGAAATA 425

QY 64 GlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIle 83  
DB 426 GGATTGATTACTCCCGAAGAGTACATTGATAGACCACCACTGCTGTAAACCGGATT 485

QY 84 GlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProAlaSer 103  
DB 486 GGTTAGCTTGATGGAGAGCTGCAAGAGGCTATAAAGTATCCTTGATGCCCTCAACT 545

QY 104 MetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAsp 123  
DB 546 ATGAGCTTAGAGAGAATCATCTGAGGCACTAGTGCAGACCTTCATCTCGGAC 605

QY 124 ProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluGluIleLeuAlaLysThrPro 143  
DB 606 CAGCGCATAGCGCTTAAAGAAATGTTGGAATACTGAAGCATTTTAAAGCAAACTCCT 665

QY 144 AsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThr 163  
DB 666 GGTGTTACATTCACAAACATTTGAATCTGCAAAACCCGAGATTGATTACGAACC 725

QY 164 ThrGlyProGluIleThrLysGlySerAspGlyLysIleAspAlaPheValSerGlyIle 183  
DB 726 ACGGACCGGAATATGAGAGATTCAAGCCGGGAAAGTAGATATATTGGTCGCGCTA 785

QY 184 GlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsnProAsnIle 203  
DB 786 GGGACTGTGGAAGTCTACTGAGTAGGGAAGTTCTCAAGAGCAGAAACAAGACATC 845

QY 204 LysLeuIleGlyValGluProValGluSerProValLeuSerGlyLysProGlyPro 223  
DB 846 AAGTTGTGTGTGGAACCAAGTAGAAAGTCCGTACTTAGCGGAGGTCAACAGCTCCA 905

QY 224 HisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeuLeu 243  
DB 906 CATTGATTCAAGGAATTGGCTCTGGTATCGTCCCATTTCAATTGGACCTTAACCATTTGT 965

QY 244 AspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeuLeuAlaLeu 263  
DB 966 GATGAATATTATTCAAGTGCAGGCTGAGAGGCTATTGAAACAGCCAGCTTCTTGCCCTC 1025

QY 264 LysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAlaPheGlnIle 283

```

Db      1026 AAAGAAGATTACTGGTGGGAATATCCTCTGAGCCGACAGCGGCTGCGTTAAAGTT 1085
QY      284  AlAlYsArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSerPheGlyGlu 303
Db      1086 GCAAAAGCGGCGGAGAAACGCGGGGAAACTCATTGGTGGTTTTCCTAGTGAAGAGAA 1145
QY      304  ArgTyrLeuSerSerValLeuPheGluSerValArgGluAlaGluSerMetThrPhe 323
Db      1146 CGTTATTATCGACTAACTGTTGATTGATTAGATATGAAGCAGAAATTGCTATT 1205
QY      324  Glu 324
Db      1206 GAA 1208

```

Search completed: May 21, 2004, 22:24:50  
 Job time : 540 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 21, 2004, 20:35:52 ; Search time 4694 Seconds  
(without alignments)  
3000.956 Million cell updates/sec

Title: US-09-931-457A-31  
Sequence: 1623  
1 MAVRSGLAKDVTLEIKTPT.....LSSVLFESVRRRAESMTFEP 325

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_plus\_p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US09931457/runat\_18052004\_121729\_20297/app\_query.fasta\_1.519  
-DB=GenEmbl -QFMT=fastapp -SUPFIX=p2n.rge -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZES=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09931457@cgn.1.1 2496@runat\_18052004\_121729\_20297 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:\*

1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_scs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_or:\*  
21: em\_ov:\*  
22: em\_pat:\*  
23: em\_ph:\*  
24: em\_pl:\*  
25: em\_ro:\*  
26: em\_sy:\*  
27: em\_un:\*  
28: em\_vi:\*  
29: em\_wt:\*

29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1622	99.9	1267	8	AF452451	AF452451 Glycine m
2	1453	89.5	1855	8	CNAPCCS7	D28777 Citrullus v
3	1451	89.4	1247	8	AY450295	AY450295 Nicotiana
4	1445	89.0	1354	8	AB029511	AB029511 Solanum t
5	1435	88.4	1308	8	AF044172	AF044172 Solanum t
6	1431	88.2	1281	8	SPIQAL	D10476 Spinach mRN
7	1431	88.2	1303	6	E06716	E06716 CDNA encodi
8	1417	87.3	1280	8	BTY10845	Y10845 Brassica ju
9	1409	86.8	1238	8	BTY10847	Y10847 Brassica ju
10	1398	86.1	1233	8	AY063827	AY063827 Arabidops
11	1384	85.3	1253	8	ATOACLY	X80376 Arabidops
12	1381	85.1	1337	8	AK071279	AK071279 Oryza sat
13	1379	85.0	1332	8	ATCYS3A	X84097 A.thaliana
14	1378	84.9	1319	8	ZMCSOATL	X85803 Z.mays mRNA
15	1377	84.8	1363	8	AF073697	AF073697 Oryza sat
16	1374	84.7	966	6	AX506198	AX506198 Sequence
17	1374	84.7	966	6	AX651825	AX651825 Sequence
18	1373	84.6	1671	8	AB040503	AB040503 Allium tu
19	1371	84.5	1355	8	BT009041	BT009041 Triticum
20	1366.5	84.2	1254	8	AY045825	AY045825 Arabidops
21	1364	84.0	1371	8	WHTCYS1	D13153 Triticum ae
22	1360	83.8	993	6	AK652865	AK652865 Sequence
23	1360	83.8	1280	8	AK061993	AK061993 Oryza sat
24	1360	83.8	1281	8	AK099598	AK099598 Oryza sat
25	1359	83.7	5259	6	AX653890	AX653890 Sequence
26	1356	83.5	1290	8	AF073695	AF073695 Oryza sat
27	1348	83.1	1192	8	ATH011976	AJ011976 Arabidops
28	1334	82.2	1562	8	AK099217	AK099217 Oryza sat
29	1317	81.1	1602	8	AK065007	AK065007 Oryza sat
30	1317	81.1	2104	8	AK065652	AK065652 Oryza sat
31	1317	81.1	1378	8	AB029512	AB029512 Solanum t
32	1316	81.1	1404	8	AF044173	AF044173 Solanum t
33	1314	81.0	1555	8	AK104704	AK104704 Oryza sat
34	1314	81.0	953	8	CAR6024	AJ006024 Cicer ari
35	1290	79.5	1377	8	ATH271727	AJ271727 Arabidops
36	1271	78.3	1460	8	AY128885	AY128885 Arabidops
37	1271	78.3	1483	6	E08016	E08016 CDNA encodi
38	1271	78.3	1484	8	SPICYSSY	D14722 Spinacia ol
39	1271	78.3	1556	8	AY099721	AY099721 Arabidops
40	1271	78.3	1492	8	AF078693	AF078693 Chlamydom
41	1270	77.7	1210	8	AY096681	AY096681 Arabidops
42	1261	77.7	1378	8	ATQACTLY2	X80377 Arabidops
43	1261	77.7	1432	8	SOCYSK	X66860 S.oleracea
44	1261	77.7	1489	8	AY065375	AY065375 Arabidops
45	1261	77.7	1489	8	AY065375	AY065375 Arabidops

ALIGNMENTS

AF452451  
LOCUS 1267 bp mRNA linear PLN 21-JAN-2002  
DEFINITION Glycine max cysteine synthase mRNA, complete cds.  
ACCESSION AF452451  
VERSION AF452451.1 GI:18252505  
KEYWORDS  
SOURCE  
ORGANISM Glycine max (soybean)  
Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

REFERENCE  
1 (bases 1 to 1267)  
AUTHORS Demosthenis, C. and Krishnan, H.B.  
TITLE Nucleotide sequence of a soybean cDNA encoding cysteine synthase  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1267)  
AUTHORS Demosthenis, C. and Krishnan, H.B.  
TITLE Direct Submission  
JOURNAL Submitted (28-NOV-2001) USDA-ARS, University of Missouri, 108W  
Curtis Hall, Columbia, MO 65211, USA  
FEATURES  
source  
Location/Qualifiers  
1..1267  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
82..1059  
/codon\_start=1  
/product="cysteine synthase"  
/protein\_id="AA166291.1"  
/db\_xref="gi:18252506"

## CDS

/translaton="MAVERSGIAKDVTEIIGKTPLVYLNKLDGCVARVNAKLELMEP  
CSSVKORIGYSMADAEKGLITPKSVLEBPTSNGTIGLAFMAARGYLITMPA  
YMSLRRIITLAFALVLTDPKEMKAVQKABEILAKTNAYILQOFENPANPKVH  
SETGPEIKRSGSDKIDAFVSGITGGITGAGKYLKCONPNKILGVEPVSPTLSG  
GKPGPHKIOGIGAGFTPGVLEVNLEDEVIQISSDEALETAKILALKEGLFVGISSGAA  
AAAFQIAKRPENAKGLIVAVFPSPGERYLSVLPESVRRASMTFEP"

## ORIGIN

## Alignment Scores:

Pred. NO.: 2,366-117 Length: 1267  
Score: 1622.00 Matches: 324  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.69% Mismatches: 0  
Query Match: 99.94% Indels: 0  
DB: Gaps: 0

US-09-931-457A-31 (1-325) x AF452451 (1-1267)

QY 1 MetAlaValGluArgSerGlyIleAlaIysAspValThrGluLeuIleGlyLysThrPro 20  
DB 82 ATGGCTGTGAAGGTCCGGAATTGCCAAGATGTTACGGAATTGATTGTAACCCCA 141  
QY 21 LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaIaArgValAlaAlaLysLeuGlu 40  
DB 142 TTAGTATATCTAAATAAATTGCGGATGTTGTGTCGCCGGTGTGCTGCTAAACTGGAG 201  
QY 41 LeuMetGluProCysSerSerValIysAspArgIleGlyTyrSerMetIleAlaAspAla 60  
DB 202 TTGATGAGCCATGCTCTAGTGTGAAGACAGAGATTGGTATAGTATGATTGCTGATGCA 261  
QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80  
DB 262 GAAGAGAGGAGCTATACACCTGGAAGAGTGTCTCATTTGAGCCCAAGTGTGTAAT 321  
QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaIaArgGlyTyrLysLeuIleIleThrMet 100  
DB 322 ACTGGCATTGATTCCTTCATGCGACAGCCAGGGGTTACAAAGCTCATTAATTACAATG 381  
QY 101 ProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120  
DB 382 CTGCTTCTATGAGTCTTGAGAGAGAAATCATTTATTAAGCTTTTGAAGCTGAGTGTGTT 441

QY 121 LeuThrAspProAlaIaLysGlyMetLysGlyAlaValGlnLysAlaGluGluIleLeuAla 140  
DB 442 TTGACAGATCTCTGCTAAGGAGATGAAGGTGCTGTTGACAGAGCTGAAGAGATATTGGCT 501  
QY 141 LysThrProAsnAlaTyrIleLeuGluGlnPheGluAsnProAlaAsnProLysValHis 160  
DB 502 AAGACGCCCAATGCTTACATTAATCTCAACAATTGAAACCCTGCCAATCCCAAGTTCAAT 561  
QY 161 TyrGluThrThrGlyProGluIleThrLysGlySerAspGlyLysIleAspAlaPheVal 180  
DB 562 TATGAACCACTGCTGTCAGAGATATGGAAGGCTCCGATGGGAAATATGATGATTTGCT 621  
QY 181 SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsn 200  
DB 622 TCTGGATAGCCACTGCTGTCATTAACAGAGTGTGCTGGAATAATATCTTAAGAGCAGAAAT 681  
QY 201 ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLys 220  
DB 682 CCGAATATTAAGCTGATTTGTTGGAACCAAGTTGAAGTCCAGTCTCTCAGAGAGAAAG 741  
QY 221 ProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240  
DB 742 CTTGGTCCACACAGATTCAAGGATTTGCTGCTGTTATCCCTGCTGTCTGGAAGTC 801  
QY 241 AsnLeuLeuAspGluValAlaGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260  
DB 802 AATCTTCTTGATGAAGATTATTAATCAATCAAGTATGAACCAATAGAACTGCAGAGCTT 861  
QY 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla 280  
DB 862 CTTCCTTTAAGAGGCTTATTTGTGGAAATATCTTCGGAGCTGCAGCTGCTGCTGCT 921  
QY 281 PheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSer 300  
DB 922 TTTCAGATTGCAAAAAGACCAAGAAATGCCGGAGGCTTATTTGCGGTTTCCAGC 981  
QY 301 PheGlyGluArgTyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSer 320  
DB 982 TTGGGAGAGAGTACTGTCCTCCGTCTATTGAGTCAATGAGACGCGAAGCTGAAGC 1041

QY 321 MetThrPheGluPro 325  
DB 1042 ATGACTTTTGAAGCCC 1056

## RESULT 2

## CNPCCS7

## LOCUS

## DEFINITION

Citrus vulgaris mRNA for cysteine synthase, complete cds.

1855 bp mRNA linear PLN 28-JAN-2003

## ACCESSION

D28777.1

GI:466530

O-acetyl-L-serine (thiol)-lyase; O-acetyl-L-serine acetate-lyase

(adding hydrogen sulfide); O-acetylserine sulfhydrylase.

## KEYWORDS

Citrus lanatus (watermelon)

## SOURCE

## ORGANISM

Citrus lanatus  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Cucurbitales; Cucurbitaceae; Citrus.

## REFERENCE

## AUTHORS

Noji, M., Murakoshi, I. and Saito, K.

Molecular cloning of a cysteine synthase cDNA from Citrus vulgaris (watermelon) by genetic complementation in an Escherichia coli Cys- auxotroph

Mol. Gen. Genet. 244 (1), 57-66 (1994)

## JOURNAL

## MEDLINE

94316193

8041362

2 (bases 1 to 1855)

Saito, K.

## TITLE

Submitted (28-FEB-1994) Kazuki Saito, Chiba University, Faculty of

Pharmaceutical Sciences, Lab. Mol. Biol. Biotech.; Yanoi-cho 1-33,

Inage-ku, Chiba, Chiba 263, Japan (E-mail: ksaito@ph.chiba-u.ac.jp,

## COMMENT

On Mar 30, 1994 this sequence version replaced gi:391655.

D17363:Submitted(05-Aug-1993)to by:  
Kazuki Saito.

## FEATURES

## source

1. 1855  
/location/Qualifiers  
/organism="Citrus limonata"  
/mol\_type="mRNA"  
/db\_xref="taxon:3654"  
/clone="pCCS7"  
/tissue\_type="etiolated seedlings"  
564. 1541  
/EC\_number="4.2.99.8"  
/codon\_start=1  
/product="cysteine synthase"  
/protein\_id="BAA05965.1"  
/db\_xref="GI:540497"  
/translation="MADAKSTAKDVTBLIGNTPLYLNRYVDCVAVAAKLEMMER  
CSYVDRIQYSMISDAENKGLTPGESVLIPTSGNTGIGLAFIAAKGYRLICMPA  
SMLEERTILRAFGAELVTDPAKMGKAVOKAEIKAKTPNSYILOQFENPAPKIH  
YETGPEIWRSGSGKIDALVSGIGTGTGAGKYLKEQNPNIKLYGVEPVESAILSG  
GKPGHKIQIGAGPIPGVLDPVNLDEVIOVSSESIETAKLILAKGLVGISSGAA  
AAAIRIAKRPENAGKLIVAVPSPFGRYLSLTVLFESVRETNMVEP"

## CDS

Alignment Scores:  
Pred. No.: 5.63e-104 Length: 1855  
Score: 1453.00 Matches: 283  
Percent Similarity: 94.46% Conservative: 24  
Best Local Similarity: 87.08% Mismatches: 18  
Query Match: 89.53% Indels: 0  
DB: 8 Gaps: 0

## ORIGIN

US-09-931-457A-31 (1-325) x CNAPCCS7 (1-1855)  
QY 1 MetAlaValGluArgSerGlyIleAlaIleAspValThrGluLeuIleGlyLysThrPro 20  
DB 564 ATGGCTGATGCAAGTCTACATAGCTAAAGATGTTACTGACTAATTGGGAATACACCA 623  
QY 21 LeuValThrLeuAsnLysLeuAlaAspGlyCysValAlaIleArgValAlaAlaLysLeuGlu 40  
DB 624 CTGTATACCTCAACCGTGTGTGATGGCTCGCTCGGCTGAGCTGCCAAGTTGGAG 683  
QY 41 LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60  
DB 684 ATGATGAGCCTTGCTCCAGTGTCAAGATAGGATTTGCTATAGCATGATTCAGATGCA 743  
QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80  
DB 744 GAAACACAGGCTTATCTACTCCGGGGAGAGATGTTCTTGAATTGAACCTTACCAAGTGTAT 803  
QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaIleArgGlyTyrLysLeuIleIleThrMet 100  
DB 804 ACTGTATAGGCTTGSCCTTCATTGCTGCTGCCAAGGTTATAGACTTATTAATTGCAATG 863  
QY 101 ProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120  
DB 864 CCTGCATCAATGATCTTGAAGAAGAACCAATCTTCGAGCTTTGGAGCTGAACCTGGIT 923  
QY 121 LeuThrAspProAlaLysGlyMetLysGlyAlaValGluLysAlaGluGluIleLeuAla 140  
DB 924 CTCACAGATCCAGCTCGGGGATGAAGAGCACTTCAAAAGGCTGAAGAGATAAAGGCA 983  
QY 141 LysThrProAsnAlaTyrIleLeuGluGlnPheGluAsnProAlaAsnProLysValHis 160  
DB 984 AAAACACCTTAATCTTACATCTTCAGCAATTGAAAACCCCTGCTAACCCGAGATCCAC 1043  
QY 161 TyrGluThrThrGlyProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheVal 180  
DB 1044 TATGAGACCACTGGTCCAGAGATCTGAGGGCTCAGGTGGAAGATGACGCACTTGTCT 1103  
QY 181 SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsn 200  
DB 1104 TCTGTATAGGACTGAGAGTACAGTACAGTACAGGAGGATATCTCAAGAACCAAAAT 1163

QY 201 ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyGlyLys 220  
DB 1164 CCTAATATTAAGTATATAGCGCTGAACCTGTGTAAGTGCAATCTATCTGAGGAAG 1223  
QY 221 ProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240  
DB 1224 CTGCGCCACATAAATCCAGGAATTTGGGCGAGGCTTCATCCCTGAGATTGTGATGTC 1283  
QY 241 AsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260  
DB 1284 AATTGCTAGATGAAGTTATTCAGTATCAAGTGAAGATCTAATTGAACCTGCAAGCTG 1343  
QY 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla 280  
DB 1344 CTTCATTTGAAGAAGGCTTACTGTTGTTATATCATCTGCGCGCGCTGCTGCA 1403  
QY 281 PheGlnIleAlaIleArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSer 300  
DB 1404 ATTAGATTTGCAAGAAGACCAAAATGCTGAAAGCTTATTGTCGGTATCCCGAGT 1463  
QY 301 PheGlyGluArgTyrLeuSerSerValLeuPheGluSerValArgGluAlaGluSer 320  
DB 1464 TTGGGGAACGATACCTCTCAACCGTGTGTTGAGTCTGTGAACGAGAGACTGAAAT 1523  
QY 321 MetThrPheGluPro 325  
DB 1524 ATGGTTTTCAGAGCCA 1538

## RESULT 3

AY450295 1247 bp mRNA linear PLN 22-NOV-2003  
LOCUS AY450295  
DEFINITION Nicotiana plumbaginifolia cysteine synthase (Csl) mRNA, partial

ACCESSION AY450295  
VERSION AY450295.1 GI:38350578

KEYWORDS  
SOURCE  
ORGANISM Nicotiana plumbaginifolia (curried-leaved tobacco)

Nicotiana plumbaginifolia  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Nicotiana.

REFERENCE  
AUTHORS Liszewska, F. and Sirko, A.  
TITLE 1 (bases 1 to 1247)  
Isolation of Nicotiana plumbaginifolia cDNAs encoding  
O-acetylserine (thiol) lyase and serine acetyltransferase

Unpublished  
2 (bases 1 to 1247)  
Liszewska, F. and Sirko, A.  
Direct Submission

JOURNAL  
REFERENCE Submitted (26-OCT-2003) Institute of Biochemistry and Biophysics,  
Polish Academy of Sciences, Pawinskiego 5A, Warsaw 02-106, Poland

TITLE  
JOURNAL  
REFERENCE  
Polish Academy of Sciences, Pawinskiego 5A, Warsaw 02-106, Poland

location/Qualifiers  
1. 1247  
/organism="Nicotiana plumbaginifolia"

source  
/mol\_type="mRNA"  
/db\_xref="taxon:4092"

gene  
/gene="Csl"  
/EC\_number="2.5.1.47"

CDS  
/note="NPCS1; O-acetylserine (thiol) lyase"

/codon\_start=1  
/product="cysteine synthase"

/protein\_id="AAR18402.1"

/db\_xref="GI:38350578"

/translation="GEKNGIADVTBLIGNTPLYLNRYVDCVAVAAKLEMMER  
CSYVDRIQYSMISDAENKGLTPGESVLIPTSGNTGIGLAFIAAKGYRLICMPA  
SMLEERTILRAFGAELVTDPAKMGKAVOKAEIKAKTPNSYILOQFENPAPKIH  
YETGPEIWRSGSGKIDALVSGIGTGTGAGKYLKEQNPNIKLYGVEPVESAILSGK  
GKPGHKIQIGAGPIPGVLDPVNLDEVIOVSSESIETAKLILAKGLVGISSGAA  
AAAIRIAKRPENAGKLIVAVPSPFGRYLSLTVLFESVRETNMVEP"

AAIRIAKRPENAGKLIVAVPSPFGRYLSLTVLFESVRETNMVEP"

## ORIGIN



Alignment Scores:  
Pred. No.: 4.83e-104 Length: 1247  
Score: 1451.00 Matches: 281  
Percent Similarity: 95.03% Conservative: 25  
Best Local Similarity: 87.27% Mismatches: 16  
Query Match: 89.40% Indels: 0  
DB: 8 Gaps: 0

US-09-931-457A-31 (1-325) x AY450295 (1-1247)

QY 4 GluArgSerGlyIleAlaIysAspValThrGluLeuIleGlyLysThrProLeuValTyr 23  
DB 4 GAAAGAGATGGAATGGCAAGATGTAAGTGAATGATCGTAACACTCTTGGTGTAC 63  
QY 24 LeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGlu 43  
DB 64 CTGAATAATGTTGTGATGGTGTGTGCGCGGTGCTGCCAAGCTGGAAGCATGGAG 123  
QY 44 ProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGluLys 63  
DB 124 CCATGCTCTAGTGTAAAGATAGATGATGTTATGATGATTAACAGATGCTGAGAGAAA 183  
QY 64 GlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyValThrGlyIle 83  
DB 184 GGCTCTGATCAAACTGGCGAGAGTGTCTCATTAACCTCAAGTGGAAACACTGGAGTA 243  
QY 84 GlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProAlaSer 103  
DB 244 GGATTGGATTATGCTGCTGCTAAAGGCTACAACTCATCATTAACGATGCTTCTTCA 303  
QY 104 MetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAsp 123  
DB 304 ATGAGCTTGAGAGAGAAATTAATCTCCGCTTTGCTGCTGAGTTGTAACCTTACCGAT 363  
QY 124 ProAlaLysGlyMetLysGlyAlaAlaGlnLysAlaGluGluIleLeuAlaLysThrPro 143  
DB 364 CCAAGCAAAAGGATGAAGGTTCTATTGAGAGCTGAAGAAATAAGGCCAAGACACACC 423  
QY 144 AsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThr 163  
DB 424 AACTCCTTATTCTTCACAATTGAAAACTGCAAAACCCAAAGGTACACTATGAGACC 483  
QY 164 ThrGlyProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheValSerGlyIle 183  
DB 484 ACTGTCCTGAGATCTGGAAGGCTCAAAATGGGAAGTAGATGCTCTGCTCTGGAATT 543  
QY 184 GlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluLysProAsnIle 203  
DB 544 GGAACAGAGAGCACATTAACAGCTTCAGGCAAGTAATTAAGAGAGACAGAACCCGACATA 603  
QY 204 LysLeuIleGlyValGluProValGluSerProValLeuSerGlyLysProGlyPro 223  
DB 604 AAGCTCTATGTTGGAACCACTGAAGTGTATTCTTCTGAGGAAGCTGTCTCA 663  
QY 224 HisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeu 243  
DB 664 CATAAGATCCAGGGGATGTTGCTGTTTCTGTTTCTGTTTGAAGTTGCTTATT 723  
QY 244 AspGluValValGlnLysSerSerAspGluAlaIleGluThrAlaLysLeuAlaLeu 263  
DB 724 GATGAAGTAATTCAGTTCAAGTGAAGAGCCATAGAAACTGCAAGCTTGGCCTTG 783  
QY 264 LysGluGlyLeuPheValGlyIleSerSerGlyValAlaAlaAlaAlaPheGlnIle 283  
DB 784 AAGGAAGTTTGTGCTGGAATATCATCTGCTCTCTGCTGCTGCGGCAATCAAACTT 843  
QY 284 AlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSerPheGlyGlu 303  
DB 844 GCTAAGCGCCCTGAAATGCTGGAGAGCTCATTTGTTTCCCAAGCTTCGGGGAG 903  
QY 304 ArgTyrLeuSerSerValLeuPheGluSerValArgGluAlaGluSerMetThrPhe 323

DB 904 CGATACTTCTCTGTCCTCTTGAATCTGTTAGACGGGAAGCAGACATGACCGTG 963  
QY 324 GluPro 325  
DB 964 GAGCCT 969  
RESULT 4  
AB029511 1354 bp mRNA linear PLN 19-JUL-2003  
LOCUS Solanum tuberosum PCS-1 mRNA for cytosolic cysteine synthase,  
DEFINITION complete cds.  
ACCESSION AB029511  
VERSION AB029511.2 GI:32997083  
KEYWORDS Solanum tuberosum (potato)  
SOURCE Solanum tuberosum  
ORGANISM Solanum tuberosum  
REFERENCE 1 Maruyama, A. and Ishizawa, K.  
AUTHORS Cytosolic Cysteine Synthase from Potato Sprout  
TITLE Published Only in Database (2001)  
REFERENCE 2 (bases 1 to 1354)  
AUTHORS Maruyama, A. and Ishizawa, K.  
TITLE Direct Submission  
JOURNAL Submitted (02-JUL-1999) Akiko Maruyama, Chiba University, Faculty  
of Pharmaceutical Sciences, Laboratory of Molecular Biology and  
Biotechnology, Research Center of Medicinal Resources, Yayoi-cho  
1-33, Inage-ku, Chiba, Chiba 263-8522, Japan  
(E-mail: amaru@p.chiba-u.ac.jp, Tel: 81-43-290-2906,  
Fax: 81-43-290-2905)  
COMMENT On Jul 19, 2003 this sequence version replaced gi:12081916.  
FEATURES  
source location/Qualifiers  
1..1354  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/strain="Dansyaku"  
/db\_xref="taxon:4113"  
/cell\_type="sprout"  
/note="tetraploid"  
1..1354  
/gene="PCS-1"  
85..1062  
/gene="PCS-1"  
/codon\_start=1  
/product="cytosolic cysteine synthase"  
/protein\_id="BAB20861.1"  
/db\_xref="GI:12081917"  
/translation="MAGEKIGIAKDVTELTGNTPLVYINNVGCVARVARLESMBP  
CSSVDRIGYSMTDAEEKGLIKPGBSVLIETPSGNTGVGLAFMAAAGYKLIITMPS  
SMSLRRIILRAFGAELVLTDPAKMKGAISKAEIKAKIPNSYILQGFENPANKIH  
YETGPEIWKSGSKVDALVSGIGTGTITGSGKYLEQNPVKLYGVEPVESALISG  
GKPEPHKIQIGAGFTFGLVNLIDVVOVSSRESISEMAKLALKEGLVGSGLAA  
AAAIKVAKRPFENAGKLIIVIFPSFGERYLSVLFETVREAEEMTVEP"  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.57e-103 Length: 1354  
Score: 1445.00 Matches: 279  
Percent Similarity: 93.85% Conservative: 26  
Best Local Similarity: 85.85% Mismatches: 20  
Query Match: 89.03% Indels: 0  
DB: 8 Gaps: 0  
US-09-931-457A-31 (1-325) x AB029511 (1-1354)  
QY 1 MetAlaValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrPro 20  
DB 85 ATGCGCGGGAAGATGGAATGGCAAGATGTAAGTGAATGATGTTGTTAACAATCTCT 144  
QY 21 LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40



```

Db      145 TTGGTATACCTGATAATGTTGGATGGCTGTGTGACAGTGTGCTGCCAGGCTGGAA 204
Oy      41  LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60
Db      205 AGCATGGAGCCATGCTCTAGTGTAAAGATAGGATCGGTATAGTATGATTACAGATGCT 264
Oy      61  GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80
Db      265 GAGGAGAGAGGCTCATCAACCCGGGAGAGTGTCTCATGSAACCTACGAGTGAAC 324
Oy      81  ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrTyrLeuIleIleThrMet 100
Db      325 ACTGGTGTAGATGGATGCTGCTGCTGCTAAAGGCTACAAACTCATCATTACGATG 384
Oy     101 ProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120
Db     385 CCTTCTCAATGAGTCTTGAGAGAGAGATTAATTCGCGTGTCTTGCGTGTGAGTGTG 444
Oy     121 LeuThrAspProAlaLysGlyMetLysGlyAlaValAlaGlnLysAlaGluGluIleLeuAla 140
Db     445 CTTACCGATCCAGCAAAAGGAGTGAAGGCTGCTATTTCAAAGGCTCAAGAGATTAAGGCC 504
Oy     141 LysThrProAsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHis 160
Db     505 AAATATACCACTCTCTATATCTTCTGAGCAATTTGAAAACCTGCTTAACCCAAAGATACAC 564
Oy     161 TyrGluThrThrGlyProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheVal 180
Db     565 TATGAGACCACTGGTCTGAGATCTGGAAGGCTCAATGGAAGTGAAGTCTCTAGTCT 624
Oy     181 SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsn 200
Db     625 TCTGGAATTGGAACAGAGGACAGATACGCTTCAAGCAAGTATTGAGAGAGAGAGAAC 684
Oy     201 ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLys 220
Db     685 CCCAAGCTTAAGCTGTATGCGGTGAACCAAGTGAAGTCTTCTGCTGGAAG 744
Oy     221 ProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240
Db     745 CCTGGTCCAGATAAGATTCAAGGGAGTGGCTGCTTCTTCTGCTGCTTGAAGT 804
Oy     241 AsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260
Db     805 AACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 864
Oy     261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyValAlaAlaAlaAlaAla 280
Db     865 CTGGCATTGAAGAGAGATGCTAGTGGAGATTTCATCTGCTGCTGCTGCTGCTGCTGCTGCT 924
Oy     281 PheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaPheProSer 300
Db     925 ATTAAGTCCCTAAGCGCCCTGAAGATGCTGGAAGCTCATGTTGTTATTTCCCAAGC 984
Oy     301 PheGlyLysGlyLysLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSer 320
Db     985 TTCGAGAGAGGATATCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1044
Oy     321 MetThrPheGluPro 325
Db     1045 ATGACTGTGAGCCT 1059

```

```

RESULT 5
AF044172 1308 bp mRNA linear PLN 06-JUL-1998
LOCUS Solanum tuberosum cysteine synthase mRNA, complete cds.
DEFINITION
ACCESSION AF044172
VERSION AF044172.1 GI:3290019
KEYWORDS Solanum tuberosum (potato)
SOURCE Solanum tuberosum
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

```

```

REFERENCE
AUTHORS Heese,H. and Hoefgen,R.
TITLE Isolation of cDNAs encoding cytosolic (Accession No. AF044172) and
JOURNAL Plant Physiol. 116, 1604 (1998)
REFERENCE
AUTHORS Heese,H.
TITLE Direct Submission
JOURNAL Submitted (22-JAN-1998) Freie Universitaet Berlin, Institut fuer
Angewandte Genetik, Albrecht-Thaer-Weg 6, Berlin 14195, Germany
FEATURES
Source
1. 1308
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Berolina"
/db_xref="taxon:4113"
35..1012
/note="CS-A, O-acetylserine (thiol) lyase; cytosolic
isoform"
/codon_start=1
/product="cysteine synthase"
/protein_id="AAC25635.1"
/db_xref="GI:3290020"
/translation="MAGEKIGIAKDVTELTGNTPLVLTNNVVDGCVARVAKLESMER
CSSVKDRIGSMITDAEKGILKPGESVLIPTSGNTGVGLAFMAAAGKLIITMPS
SMLEKRIILRGSESLVLTDPKGMKGAISKAESIKAKTPNSYILQEPANPKIH
YETTPRIWKSNGKVDALASGIGTGTGSGKYLREONPNVKLYGVSEYSAIISG
GKPGHKIGIGAGFIPGVLEVNLDVGVSSSESIEMAKLALKKGLVGISSGAA
AAAAIKVAKRPENAKLIVIPSPFGRYLSVLFETVREAENMTVEP"
ORIGIN
Alignment Scores:
Pred. No.: 9.05e-103 Length: 1308
Score: 1435.00 Matches: 277
Percent Similarity: 93.23% Conservative: 26
Best Local Similarity: 85.23% Mismatches: 22
Query Match: 88.42% Indels: 0
DB: 8 Gaps: 0
US-09-931-457A-31 (1-325) x AF044172 (1-1308)
Oy      1  MetaLalaValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrPro 20
Db      35  ATGGCGGGGGGAAAGATTGGATTGCCAAGATGATGATGATGATGATGATGATGATGATGAT 94
Oy     21  LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40
Db     95  TTGGTATACCTGATATATGTTGATGATGGTGTGTGACGATGTTGCTGCCAAGCTGGA 154
Oy     41  LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60
Db    155  AGCATGGAGCCATGCTCTAGTGTAAAGATAGGATGCGGTATAGTATGATTACAGATGCT 214
Oy     61  GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80
Db    215  GAGGAGAGAGGCTCATCAACCCGGGAGAGTGTCTTCATCGAAGCTACGAGTGAAC 274
Oy     81  ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMet 100
Db    275  ACCGGTGTAGATGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 334
Oy    101  ProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120
Db    335  CCTTCTCAATGAGTCTTGAGAGAGAGATTAATTCGCGTGTCTTGGTCTGAGTGTAGTG 394
Oy    121  LeuThrAspProAlaLysGlyMetLysGlyAlaValAlaGlnLysAlaGluGluIleLeuAla 140
Db    395  CTTACCGATCCAGCAAAAGGAGTGAAGGCTGCTATTTCAAAGGCTGAAGAGATTAAGGCC 454
Oy    141  LysThrProAsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHis 160

```

Db 455 AAAACACCACTCCTATATCTTTCAGCAATTTGAAAACCTGCTAACCCAAAGATACAC 514

QY 161 TyrGluThrGlyProGluIleThrPlysGlySerAspGlyLysIleAspAlaPheVal 180

Db 515 TATGAGACCACTGCTCTGAGATCTGAAAGGCTCAACGGAAAGTGAATGCTTAGCC 574

QY 181 SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrIleuLysGluGlnAsn 200

Db 575 TCTGAATTGGAACAGGAGGACGATTAAGTTCAGGCAAGTATTGAGAGCAGAAC 634

QY 201 ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLys 220

Db 635 CCCAACGTTAAGCTGATGCGCTGGAACCAAGTGAAGTCTATCCTTCTGCTGGAAG 694

QY 221 ProGlyProHisLysIleGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240

Db 695 CTTGTCACATAGATTCAAGGCAATGCTGCTGCTTCAITTCCTGCTGTTTGAAGTT 754

QY 241 AsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260

Db 755 AACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 814

QY 261 LeuAlaLeuLysGlyGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla 280

Db 815 CTGGCAATGAGAGGAGGATGCTAGTGGGAATTTCACTGCTGCTGCTGCTGCTGCTGCT 874

QY 281 PheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaAlaPheProSer 300

Db 875 ATTAAGTGTAGTAAAGCGCTTGAAGTCTGAGGAGCTCAATGTTATTTTCCCAAGC 934

QY 301 PheGlyGluArgTyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSer 320

Db 935 TTCGAGAGCGATATCTTCTCTGCTACTCTTCCAACTGTCAAGAGAGAGAGAGAGAAC 994

QY 321 MetThrPheGluPro 325

Db 995 ATGACTGTGAGCCT 1009

RESULT 6

SPIOTAL

LOCUS Spinach mRNA for O-acetylserine(thiol)-lyase, complete cds. 1281 bp mRNA linear PLN 02-FEB-1999

DEFINITION D10476 GI:218276

ACCESSION D10476

VERSION D10476.1

KEYWORDS O-acetyl-L-serine acetate-lyase (adding hydrogen sulfide); O-acetylserine(thiol) lyase; cysteine synthase.

SOURCE Spinacia oleracea (spinach)

ORGANISM Spinacia oleracea

REFERENCE 1. Saito, K., Miura, N., Yamazaki, M., Hirano, H. and Murakoshi, I. Carboxylates; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Amaranthaceae; Spinacia.

AUTHORS Saito, K., Miura, N., Yamazaki, M., Hirano, H. and Murakoshi, I.

TITLE Molecular cloning and bacterial expression of cDNA encoding a plant cysteine synthase

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 89 (17), 8078-8082 (1992)

MEDLINE 92390392

PUBMED 1518833

REFERENCE 2 (bases 1 to 1281)

AUTHORS Saito, K.

TITLE Direct Submission

JOURNAL Submitted (12-FEB-1992) Kazuki Saito, Chiba University, Faculty of Pharmaceutical Sciences, Lab. Mol. Biol. Biotech.; Yayoi-cho 1-33, Inage-ku, Chiba, Chiba 263, Japan (E-mail:ksaito@chiba-u.ac.jp, Tel:043-290-2904, Fax:043-290-2905)

COMMENT Submitted (12-FEB-1992) to DDBJ by: Faculty of Pharmaceutical Sciences, Chiba University Yayoi-cho 1-33 Chiba 260 Japan Phone: 0472-51-1111 x2707 Fax: 0472-55-1574. Kazuki Saito

FEATURES

Source

Location/Qualifiers

1..1281

/organism="Spinacia oleracea"

/mol\_type="mRNA"

/db\_xref="taxon:3562"

52..1029

/EC\_number="4.2.99.8"

/codon\_start=1

/product="O-acetylserine(thiol) lyase"

/protein\_id="BA01279.1"

/db\_xref="GI:218276"

/translation="MTEKAFIAKDVTELIGKPLVLTNVAAGCVARVAALKEGMEP CSSVKDRIGFSMTDAEKSGLITPESVLIETPSGNTGIGLAFIAAGYKLIITMFA SMLERRTILRAFGAHLITDPKAKMGAVOKAERTDKTPNSYILOQFENPANPKYH YETTPKIKWGTGKIDIFVSGITGCTITGAGKYLEONPDVKLIGLEPVEASAVLSG GRPDPKIQGLAGFIPGLVDVNIIDEVVOISSESEIEMAKLALKKGLVIGISSGAA AAAIKVAKRPENAKLIVAVFPFGBRYLSSVLPDSVFKAESMVIES"

1281

polya\_site

ORIGIN

Alignment Scores:

Pred. No.: 1,81e-102 Length: 1281

Score: 1431.00 Matches: 281

Percent Similarity: 94.14% Conservative: 24

Best Local Similarity: 86.73% Mismatches: 19

Query Match: 88.17% Indels: 0

Gaps: 0

US-09-931-457A-31 (1-325) x SPIOTAL (1-1281)

QY 1 MetAlaValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrPro 20

Db 52 ATGTTGAG 111

QY 21 LeuValIleLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40

Db 112 TTGGTATATCTCAACACTGTCGCGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 171

QY 41 LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60

Db 172 GGAATGAACTTGTCTAGTGTAAAGACAGAGATGGGTTCAATGATTAATGATGATGCT 231

QY 61 GluGlySerGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80

Db 232 GAAAAAGCGGCTTATTACACCTGAGAGAGTGTCTGATTGAGCCCAAGTGGAAT 291

QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMet 100

Db 292 ACTGCATGATTAAGCTTCACTTCACTGAGAGAGCTAAAGGTTACAGCTCAATTAAGATG 351

QY 101 ProAlaSerMetSerLeuGluArgArgIleIleLeuAlaAlaPheGlyAlaGluLeuVal 120

Db 352 CCAGCATCAATGAGCTTGAAGGAGAGCTAATTTCAAGGCGCTTGTGCTGAGCTTATC 411

QY 121 LeuThrAspProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluGluIleLeuAla 140

Db 412 CTTACTGATCAAGCAAAAGGTATGAAGGGGCTGTTCAAGAGCTGAGAGATCCGTGAC 471

QY 141 LysThrProAsnAlaTyrIleLeuGlnGluPheGluAsnProAlaAsnProLysValHis 160

Db 472 AAAACTCCTAATTCATATATATACATCAACAGTTGAAAACCTGCAACCCAAAGGTTCA 531

QY 161 TyrGluThrThrGlyProGluIleThrPlysGlySerAspGlyLysIleAspAlaPheVal 180

Db 532 TATGAACAACCTGACACAGAAATTTGAAAAGCACAGGTGAAATAATGATATATTCGTC 591

QY 181 SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrIleuLysGluGlnAsn 200

Db 592 TCTGAATAGGAGCTGAGGTACATTAACAGGTGCAAGAAATACCTAAAGGAACAAAC 651

QY 201 ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLys 220

Db 652 CCGATGTTAAGCTAATGCTGGAACCAAGTGAAGAGTGTGTTATGCTGAGAGAGAAA 711

Qy	221	ProGlyProHisLysIleGlnGlyIleGlyValaGlyPheIleProGlyValIleuGluVal	240
		::::	::::
Db	712	CCTGCCCAACATAAGATTCAAGCACTTGAGCTGGATTCACTGCTGTTCTGAGATGTG	771
Qy	241	AsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysIleu	260
		::::	::::
Db	772	AATATTATCATGAAGCTGGTTCAAGATATCCAGTGAAGAATCTATTGAAATGCCAAATTG	831
Qy	261	LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla	280
Db	832	CTGCCCTCAAGGAAGGTTACTGCTGGATTTCATCTGGTCTGCTGCTGCCGCTGCC	891
Qy	281	PheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValaPheProSer	300
		::::	::::
Db	892	ATTAAAGTCGCAAGAGGCTTCAAAATGCTGGAAAATCATGCTGCTGCTTTCCAGC	951
Qy	301	PheGlyGluArgTyrIleuSerSerValIleuPheGluSerValaArgGluAlaGluSer	320
		::::	::::
Db	952	TTTGGCGAAGCATATTATCTCTCGGTGTTGTTTGATTCAAGTGAAGAGGACAGAGAGC	1011
Qy	321	MetThrPheGlu	324
Db	1012	ATGGTTATTGAG	1023

RESULT	7
E06716	
LOCUS	1303 bp RNA linear PAT 29-SEP-1997
DEFINITION	cDNA encoding Spinacia cysteine synthetase.
ACCESSION	E06716
VERSION	E06716.1 GI:2174898
KEYWORDS	JP 1994038770-A/1.
SOURCE	Spinacia oleracea (spinach)
ORGANISM	Spinacia oleracea Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Amaranthaceae; Spinacia. 1 (bases 1 to 1303) Saitou, K. and Murakoshi, I. GENE CODING CYSTEINE SYNTHETASE Patent: JP 1994038770-A 1 15-FEB-1994; MITSUBISHI CORP, MITSUBISHI KASEI CORP OS Spinacia oleracea
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

PF	05-FEB-1992 JP 1992020315	
PI	SAITOU KAZUSUE, MURAKOSHI ISAMU	
PC	C12N15/60, C12N9/88, (C12N9/88, C12R1:19);	
CC	strandedness: Double;	
CC	topology: Linear;	
CC	hypothetical: No;	
CC	anti-sense: No;	
CC	*source: strain=parade;	
CC	*source: clone=ptv118N;	
FH	Key	Location/Qualifiers
FH		
FT	5'UTR	1..51
FT	CDS	52..1029
FT		/product='Spinacia cysteine synthetase' FT
3'UTR		1030..1303.

```

FEATURES
    source
        location/Qualifiers
            1..1303
            /organism="Spinacia oleracea"
            /mol_type="genomic RNA"
            /db_xref="taxon:3562"

```

Alignment Scores:	
Pred. No.:	1,85e-102
Score:	1431.00
Percent Similarity:	94.14%
Best Local Similarity:	86.73%
Query Match:	88.17%
Length:	1303
Matches:	281
Conservative:	24
Mismatches:	19
Indels:	0

```
DB:                               6                                Gaps:                               0
US-09-931-457A-31 (1-325) x E06716 (1-1303)

QY      1 MetalavalGluArgSerglyllealalaAspValThrGluLeuIleGlyIysThrPro 20
        |||          ||::||:|||||
Db       52 ATGGTTAGAGAGAAAGCCCTTCATTGCTAAAGATGTCACTGAATTGATTGGAAAACGCCA 111

QY      21 LeuValTyrLeuAsnLysLeuAlaaspGlyCysValAlaArgValAlaAlaLysLeuGlu 40
        |||          |||::|||
Db       112 TTGGTATATCTCAAACACTGTGCACCGATGGTGTTGTGCTCGTGTCTGCAAAAGCTGAA 171

QY      41 LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60
        |||          |||::|||
Db       172 GGATGGAAACCTTGCTCTAAGGTTAAGACACAGCATTTGGGTTCAATAFATTACTGAGACT 231

QY      61 GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerglyAsn 80
        ||::|||          |||::|||
Db       232 GAAAAAAGCGGGCTTATTACAACCTGAGAGAGAGTGTCTGATTGAGCCCAACAGTGAAT 291

QY      81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrIlysLeuIleIleThrMet 100
        |||          |||::|||
Db       292 ACTGCATTGGAATAGACCTTCATCGCACGACGCTAAAGTTACAAGCTCATCATTAGATG 351

QY      101 ProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120
        |||          |||::|||
Db       352 CCAGCATCAATGAGTCTTGAGCGGAGACGACTATTCTCAGGGCCTTTGGTGCTGAGCTTATC 411
```

QY	121	LeuThrAspProAlaIalysGlyMetIysGlyAlaValGlnLysAlaGluGluIleLeuAla	140
Db	412	CTTACTGATCCAGCAAAAGGTATGAAAGGGCTGTTCAGAAAGGCTGAGAGATCCGTGAC	471
QY	141	LYSThrProAsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHis	160
Db	472	AAAACCTCCTAATTCATATATATACACACAGATTGAAAACCCCTGCCAACCCAAAGGTTCAAT	531
QY	161	TyrGluThrThrGlyProGluIleTyrPlySgIySeraSpsGlyLysIleAspAlaPheVal	180
Db	532	TATGAACAACACTGCACCGAATTTGGAAAGGCGACACGTGAAAAAATTGATATATATTCGTC	591
QY	181	SerGlyIleGlyIThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsn	200
Db	592	TCTGAATAGGCACTGGAGGTACATAAACAGGTGCAGAAAATATCCTTAAAGGAACAAAC	651

Db	652	CCGATGTTAAGCTAATTGGCCCTGGAACCAAGTGAAGTCTGTATTGTCTGGAGGAAA	711
Qy	221	ProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal	240
Db	712	CCTGGCCCACTAAGATTCAGAAGACTTGAAGCTGCATTCATACCTGCTGTTCTGGATGTG	771
Qy	241	AsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu	260
Db	772	AATATTATCGATGAAGTGGTTCAGATATCCAGTCAAGATCTATTGAATGGCCAAATTG	831
Qy	261	IeuAlaLeuLysGlnGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla	280
Db	832	CTGCGCCCTCAAGGAAGTCTACTGTTGGGATTCATCTGGTGCTGCTGCGCGCTGCC	891

Qy 281 pheglnl leala l y s a r p r o g i u s m a l a c l y l s b e u l e v a l a l a v a l e n e r u s e r 500  
 : : : : :  
 Db 892 ATTAAGTGCACAAGAGGCTGAAAATGCTGGAAACTCATCGTCGCTGTCTTTCCAGC 951  
 Qy 301 PheGlyGluArgTyrLeuSerSerValLeuPheGluSerValArgArgGluValaGluSer 320  
 : : : : :  
 : : : : :

```

QY      321 MetThrPheGlu 324
          |||
          |||
DB      1012 ATGGTATTGAG 1023

```



BUY10845

LOCUS BUY10845 1280 bp mRNA linear PLN 02-FEB-1999  
DEFINITION Brassica juncea mRNA for O-acetylserine(thiol) lyase, clone OAS-TL4.

ACCESSION Y10845  
VERSION Y10845.1 GI:2243119  
KEYWORDS O-acetylserine(thiol) lyase.  
SOURCE Brassica juncea

ORGANISM

Brassica juncea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE

1 Schaefer, H.J., Haag-Kerwer, A. and Rausch, T.  
CDNA cloning and expression analysis of genes encoding GSH synthesis in roots of the heavy-metal accumulator Brassica juncea L.: evidence for Cd-induction of a putative mitochondrial gamma-glutamylcysteine synthetase isoform  
Plant Mol. Biol. 37 (1), 87-97 (1998)  
98281577  
MEDLINE  
PUBMED 9620267  
2 (bases 1 to 1280)  
Schaefer, H.J.  
Direct Submission  
Submitted (28-JAN-1997) H.J. Schaefer, Universitaet Heidelberg, Botanisches Institut, Im Neuenheimer Feld 360, 69120 Heidelberg, FRG

FEATURES

source 1..1280  
/organism="Brassica juncea"  
/mol\_type="mRNA"  
/sub\_species="Vittasso"  
/db\_xref="taxon:3707"  
/clone="OAS-TL4"  
/tissue\_type="root"  
/clone\_id="lambda ZAP Express"  
/dev\_stage="mature"  
74..1042  
/codon\_start=1  
/product="O-acetylserine(thiol) lyase"  
/protein\_id="CAA71798.1"  
/db\_xref="GI:2243120"  
/db\_xref="GOA:O23733"  
/db\_xref="SWISS-PROT:O23733"  
/translation="MASRIADQVTEHIGNTPLVYLNVAEGCVRAAKLEMMERCCSSVKDRIGFSMISDAEQKGLIKPESVLIPTSGNTGVGLAFTAAAGYKLIITPDASMSVERITLLAFGEVLVITDPKAKKGAIAKABILAKTPNGVYLQFENPANKIHYETGPETIKGTGDKIDFVSGIGTGTITGAGKYLKEONPVKLYGVEPIESAILSGKPGPHKIGIGAGFIPSVLEVDLIDVYQVSSDSIDMARLLAKGLLVGISGAAAAAATKAKRPENAGKLFVAVPFSEGERYLTSLFDATRKASMTFOA"

CDS

Alignment Scores:  
Pred. No.: 2.22e-101 Length: 1280  
Score: 1417.00 Matches: 271  
Percent Similarity: 94.36% Conservative: 30  
Best Local Similarity: 84.95% Mismatches: 18  
Query Match: 87.31% Indels: 0  
Gaps: 0  
DB: 8

ORIGIN

US-09-931-457A-31 (1-325) x BUY10845 (1-1280)  
OY 6 SerGlyTlleAlaAlaAspValThrGluLeuIleGlyLysThrProLeuValTyrLeuAsn 25  
DB 80 TCTCGAATTCCTTAAGATGTTACCGAATTCATTGGGAACACTCCATTGGTGTATCTGAAC 139  
OY 26 LysLeuAlaAspGlyCysValAlaAlaArgValAlaAlaLysLeuGluLeuMetGluProCys 45  
DB 140 AATGTGCTGAAGATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 199  
OY 46 SerSerValLysAspArgGlyIleGlyTyrSerMetIleAlaAspAlaGluGluLysGlyLeu 65  
DB 200 TCCAGTGTCAAGACAGGATGTTTATGATGATCTCTGATGCAGAGCAAAAGGCTCTG 259

OY 66 IleThrProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIleGlyLeu 85  
DB 260 ATCAACCAAGAGAGAGAGTGTCTGATTGAGCCCAAGAGTGAACACCGAGTGGCTTA 319  
OY 86 AlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProAlaSerMetSer 105  
DB 320 GCGTTCAACGCGCGCTGCCAAGGGGTACCAAGCTCATCTATGCTGACGCTTCAATGAGT 379  
OY 106 LeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspProAla 125  
DB 380 GTGAGAGAAAGATCATCTCTTACTTTTGGAGTTGAGTTGCTTAACTGACCTGCT 439  
OY 126 LysGlyMetLysGlyAlaValAlaGlnLysAlaGluGluIleLeuAlaLysThrProAla 145  
DB 440 AAGGGATGAAAGCGCTATTCGCAAGGCCGAGAGATTTTAGCCAAACACCCCAATGCT 499  
OY 146 TyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThrGly 165  
DB 500 TACATGCTCCACACAGTTTGAAGAACCTTGCCAAACCAAGATACACTATGAGACTACTGA 559  
OY 166 ProGluIleTyrLysGlySerAspGlyLysIleAlaAspAlaPheValSerGlyIleGlyThr 185  
DB 560 CCGATGATATGAAAGGACATGATGAAATAATCATGATGATTTGTTCTGGGATGGTACT 619  
OY 186 GlyGlyTyrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsnProAlaIleLysLeu 205  
DB 620 GGTGTACCATCACAGGTGACAGGAAGTATCTCAAGAACAGAACCTTAACGTTAAGCTG 679  
OY 206 IleGlyValGluProValGluSerProValLeuSerGlyGlyLysProGlyProHisLys 225  
DB 680 TATGAGATGAGCAATGAAGTGTATCTATCCGTTGGGAAGCCAGGCTTCAACAAG 739  
OY 226 IleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeuLeuAspGlu 245  
DB 740 ATTCAAGGATGTGTGCTGCTTATACCAAGTGTGTGAGGTTGATCTTATGATGAA 799  
OY 246 ValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeuLeuAlaLeuLysGlu 265  
DB 800 GTTGTCAAGTTTCAAGTATGAATCCATGCACATGCAAGGCTTCTGCTTAAAGAA 859  
OY 266 GlyLeuPheValGlyLysSerSerGlyAlaAlaAlaAlaAlaPheGlnIleAlaLys 285  
DB 860 GGTCTTCTTGTGGGATATCATCTGTGACAGCAGCGCCGACCAATCAAACTTCAAG 919  
OY 286 ArgProGluAsnAlaGlyLysLeuIleValAlaAlaPheProSerPheGlyGluArgTyr 305  
DB 920 AGGCCAGAAATATCTGGGAAGCTTTTGTGCGCGGTCTCCGAGTTTCGAGAGAGGTAC 979  
OY 306 LeuSerSerValLeuPheGluSerValAlaArgGluAlaGluSerMetThrPheGlu 324  
DB 980 CTGTGCACTGATCTATTCATGATGCCACAGAAAGAGAGATCAATGACCTTCCAG 1036

RESULT 9  
BUY10847 1238 bp mRNA linear PLN 01-FEB-1999  
LOCUS BUY10847  
DEFINITION Brassica juncea mRNA for O-acetylserine(thiol) lyase, clone OAS-TL6.  
ACCESSION Y10847  
VERSION Y10847.1 GI:2243123  
KEYWORDS O-acetylserine(thiol) lyase.  
SOURCE Brassica juncea  
ORGANISM Brassica juncea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
REFERENCE 1 Schaefer, H.J., Haag-Kerwer, A. and Rausch, T.  
CDNA cloning and expression analysis of genes encoding GSH synthesis in roots of the heavy-metal accumulator Brassica juncea L.: evidence for Cd-induction of a putative mitochondrial gamma-glutamylcysteine synthetase isoform  
Plant Mol. Biol. 37 (1), 87-97 (1998)  
JOURNAL



MEDLINE 98281577  
PUBMED 9620267  
REFERENCE 2 (bases 1 to 1238)  
AUTHORS Schaefer, H.J.  
TITLE Direct Submission  
JOURNAL Submitted (28-JAN-1997) H.J. Schaefer, Universitaet Heidelberg,  
Botanisches Institut, Im Neuenheimer Feld 360, 69120 Heidelberg,  
FRG

FEATURES  
Source Location/Qualifiers  
1..1238  
/organism="Brassica juncea"  
/mol\_type="mRNA"  
/sub\_species="Vittasso"  
/db\_xref="taxon:3707"  
/clone="OAS-TL6"  
/tissue\_type="root"  
/clone\_lib="lambda ZAP Express"  
/dev\_stage="mature"  
18..992  
/codon\_start=1  
/product="O-acetylserine (thiol) lyase"  
/protein\_id="CAA71800.1"  
/db\_xref="GI:2243124"  
/db\_xref="GOA:O23735"  
/db\_xref="SWISS-PROT:O23735"  
/translation="MASRIGIANDVTLEIGNTPLYLNSVAEGCVGRVAKLEMBPC  
SSVKDRIGFSMTSDAEKGLIKGESVILIEPTSGNTGVLAFTAAKGYKLITMPAS  
MSIERRIILAFVELVLTDPKAKMGAIARAEILAKTPNGYMLQGENPANKIHY  
ETTGPEIWKGTGKIDGFIISGIGTGTITGAGKYLKEQNPVKLYGEPVESAILSGG  
KPGPHKIQIGAGFIPNVLNTLIDVAVQVSSDESIDMARLLAREEGLVGISSGAA  
AAAIKIAKRPENAGKLFVAVFSPFGERYLTVLFDATRKEASMTPEA"

CDS

ORIGIN

Alignment Scores:

Pred. No.: 8.93e-101 Length: 1238  
Score: 1409.00 Matches: 269  
Percent Similarity: 93.44% Conservative: 30  
Best Local Similarity: 84.06% Mismatches: 21  
Query Match: 86.81% Indels: 0  
DB: 8 Gaps: 0

US-09-931-457A-31 (1-325) x BUI10847 (1-1238)

QY 5 ArgSerGlyIleAlaIysAspValThrGluLeuIleGlyLysThrProLeuValTyrLeu 24  
DB 27 CGTATTGATTGCTAATGATGACCGAATTGATTGGACACATCCGTTGGTATCTG 86  
QY 25 AsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluPro 44  
DB 87 AACAGTGTGCTGAAGGATGTTGTCGCGCTGCCCTTAAGCTTGAGATGATGAACCC 146  
QY 45 CysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGluLysGly 64  
DB 147 TGCTCCAGCGCTCAAGACGAGATTGTTTACATGATCTGATGACAGAGAAAGGGT 206  
QY 65 LeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIleGly 84  
DB 207 CTGATTAAACCGAGAGAGTGTGTTGATTGAGCCACAGAGTGGAACACGAGATTGGG 266  
QY 85 LeuAlaPheMetAlaAlaIaargGlyTyrLysLeuIleIleThrMetProLaserMet 104  
DB 267 TTGGCATTCACGGCGGTGCCAAGGGCTACAAAGCTTATTATACCATGCGCATCAATG 326  
QY 105 SerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspPro 124  
DB 327 AGTATTGAGAGAGAAATCACTCTTACGCTTTGTTGAGTGTGTTTAACTGACCCG 386  
QY 125 AlaLysGlyMetLysGlyAlaValGlnLysAlaGluGluIleLeuAlaTyrThrProAsn 144  
DB 387 GCCAAGGCGATGAAGCGCTATCGCAAGCGGAGAGATTTTGGCCAAACACCCAAAT 446  
QY 145 AlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThr 164

DB 447 GGTACATGCTCCACACAGTTTGAGAACCTGTCTAACCCAAAGATCCACTATGAGACTACT 506  
QY 165 GlyProGluIleTyrPlyGlySerAspGlyLysIleAspAlaPheValSerGlyIleGly 184  
DB 507 GGACCTGAGATTGGAAGGACACTGAGGGTAAATCGATGCTTATTCTTGGGATTGGT 566  
QY 185 ThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluLysAsnProAsnIleLys 204  
DB 567 ACTGTGTGATACATACAGAGTGTGGAAGTATCTCAAGAACAGAACCGAAGCTTAAG 626  
QY 205 LeuIleGlyValGluProValGluSerProValLeuSerGlyGlyLysProGlyProHis 224  
DB 627 CTGTATGTTGTGAGCCAGTCCGAAAGCGCTATTCTATCTGTGGAAGCCAGCCCTCAC 686  
QY 225 LysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeuLeuAsp 244  
DB 687 AAGATTCAAGGAGTAGTGTCTGTTTATACCAATGATATTGAGACTATCTTATTGAT 746  
QY 245 GluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeuLeuAlaLys 264  
DB 747 GAAGTGTTCAGATTTCAGTATGATGATGATCCATCGACATGCGAAGGCTTCTCGCGAA 806  
QY 265 GluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAlaPheGlnIleAla 284  
DB 807 GAAGTCTTCTTGTGGGAATATCATCCGTCGACAGCACTGACAGCAATCAAACTTGA 866  
QY 285 LysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSerPheGlyGluArg 304  
DB 867 AAGAGGCCAGAAACCGCCGGAAGCTTTTCGTGCGAGTGTCCCAAGTTTCGAGAGAGG 926  
QY 305 TyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSerMetThrPheGlu 324  
DB 927 TACCTGTGACCGTACTGTTCGATGCCACAGAAAGAGAGAAATCCATGACCTTCGAG 986

RESULT 10  
AY063827 1233 bp mRNA linear PLN 23-SEP-2002  
LOCUS Arabidopsis thaliana At3g22460 mRNA sequence.  
DEFINITION Arabidopsis thaliana  
ACCESSION AY063827  
VERSION AY063827.1 GI:17380705  
KEYWORDS FLI CDNA.  
SOURCE Arabidopsis thaliana (chale cress)  
ORGANISM Arabidopsis thaliana  
REFERENCE 1 (bases 1 to 1233)  
AUTHORS Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Tripp, M.G., Wu, T., Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
TITLE Arabidopsis Full length cDNA clones  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1233)  
AUTHORS Yamada, K., Banh, J., Banno, F., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
TITLE Submitted (19-NOV-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA  
JOURNAL Submitted (19-NOV-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA  
REFERENCE 3 (bases 1 to 1233)  
AUTHORS Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinzaki, K., Davis, R.W., Ecker, J.R., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J.,

COMMENT  
Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Tripp, M.G., Wu, T., Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
Submitted (23-SEP-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA  
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamaya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinzaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, B.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shim, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.  
Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinzaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.  
Location/Qualifiers  
1. 1233

FEATURES

source  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="3"  
/clone="RAFL06-87-B19 (R11881)"  
/ecotype="Columbia"  
/note="This clone is in a modified pBluescript vector (Lambda ZAP) as a XhoI/SacI insert."  
1. 1233  
/gene="At3g22460"  
84. 1061  
/gene="At3g22460"  
/note="This may be a potential product of a partially processed transcript corresponding to gene At3g22460."  
/evidence="experimental"  
misc\_difference 1216  
/gene="At3g22460"  
/note="not present in genomic sequence"

ORIGIN

Alignment Scores:  
Pred. No.: 6.39e-100 Length: 1233  
Score: 1398.00 Matches: 270  
Percent Similarity: 92.00% Conservative: 29  
Best Local Similarity: 83.08% Mismatches: 26  
Query Match: 86.14% Indels: 0  
DB: 8 Gaps: 0

US-09-931-457a-31 (1-325) x AY063827 (1-1233)

QY 1 MetAlaValGluArgSerGlyLeuAlaValAspValThrGluLeuIleGlyIleThrPro 20  
DB 84 ATGGCTTCAGTGGCTCCCTAAATCGCTAAAGATGTAAGTAACTAATAGGGAACATCCCA 143  
QY 21 LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaAargValAlaAlaLysLeuGlu 40  
DB 144 TTAGTGTATCTAAACAAAGTAAAGATGTGTGTGTCATGTTGCTGCTAAGCTTGAG 203  
QY 41 LeuMetGluProCysSerSerValIleAspArgIleGlyTyrSerMetIleAlaAspAla 60  
DB 204 ATGATGAACACATGTTCTAGCGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 263  
QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80  
DB 264 GAAGCTAAAGACTTATTAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 323  
QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAargGlyTyrLysLeuIleIleThrMet 100

DB 324 ACTGAGTTGTTAGCATTTAAGCTGAGCTGCAAGGCTAATAAGCTTATTAACAATG 383  
QY 101 ProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120  
DB 384 CCAAGCTTCAATGATATGAG 443  
QY 121 LeuThrAspProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluGluIleLeuAla 140  
DB 444 CTTACTGATCCGCTAAAGTATGAAAGGTGCTGTTCAAGGCGGAGAGAGATTTGGCG 503  
QY 141 LysThrProAsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHis 160  
DB 504 AAACACCTAATAGTATATGCTTCAACAGTTTGAAGATTCCTAATCCAAAGATCCAT 563  
QY 161 TyrGluThrThrGlyProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheVal 180  
DB 564 TATGAGACTACTGACCTGAGATATGAAAGGTTCGTGGAAGAGAGAGAGAGAGAGAGAG 623  
QY 181 SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsn 200  
DB 624 TCTGCTATGTTGTTAGTGTGTAAGTGAACAGTGTGGAAGTATCTCAAGGAACAGAAC 683  
QY 201 ProAsnLysLeuIleGlyValGluProValGluSerProValLeuSerGlyIleLys 220  
DB 684 ACAACATAAAGCTGATAGTGTGCAAGCTGTGAAGCCCTATTCTGTCGGTGAAG 743  
QY 221 ProGlyProHisLysIleGlnIleLysSerSerAspGluAlaIleGluThrAlaLysLeu 240  
DB 744 CCAGGTCCCAACAAATTCAGGATATAGTGTGCTGTTTATTCAGGCAATTTGGATGT 803  
QY 241 AsnLeuLeuAspGluValAlaGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260  
DB 804 GATCTTATGATGATGAGTGTTCAGGTTTCAAGTGAAGATCCATGACATGCAAGGCTT 863  
QY 261 LeuAlaLeuLysGluGlyLeuPheValGlyLysSerSerGlyAlaAlaAlaAlaAla 280  
DB 864 CTGCTAGGAGAGAGAGGCTGTGTTAGTGGATCTCATCTGAGCAAGCACTACTGAGCA 923  
QY 281 PheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaAlaPheProSer 300  
DB 924 ATCAACTTGCAGAGAGGCGCAGAAATCTCTGGAGAGCTTATGAGGCGGTTCAGAT 983  
QY 301 PheGlyGluArgTyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSer 320  
DB 984 TTGGGGAGAGATATCTATCTGACTGTACTGTTTATGATGAGCAAGAGAGAGAGACT 1043  
QY 321 MetThrPheGluPro 325  
DB 1044 ATGACCTTTGAGCCC 1058

RESULT 11  
ATOACLY 1253 bp mRNA linear PLN 15-FEB-2000  
LOCUS Arabidopsis thaliana mRNA for O-acetylserine lyase (At.OAS.5-8).  
DEFINITION X80376  
VERSION X80376.2 GI:6983573  
KEYWORDS O-acetylserine(thiol) lyase.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1  
AUTHORS Hell, R., Bork, C., Bogdanova, N., Frolov, I. and Hauschild, R.  
TITLES Isolation and characterization of two cDNAs encoding for compartment specific isoforms of O-acetylserine (thiol) lyase from Arabidopsis thaliana  
JOURNAL FEBS Lett. 351 (2), 257-262 (1994)  
MEDLINE 94364481  
PUBMED 8082776  
REFERENCES 2  
AUTHORS Hell, R.  
TITLES Direct Submission

JOURNAL Submitted (18-JUL-1994) R. Hell, Ruhr-Universitaet Bochum, Lehrstuhl fuer Pflanzenphysiologie, Universitaetstr. 150, 44801 Bochum, FRG

REMARK Revised by (4)

REFERENCE 3 (bases 1 to 1234)

AUTHORS Hell, R.

TITLE Direct Submission

JOURNAL Submitted (14-DEC-1994) R. Hell, Ruhr-Universitaet Bochum, Lehrstuhl fuer Pflanzenphysiologie, Universitaetstr. 150, 44801 Bochum, FRG

REFERENCE 4 (bases 1 to 1253)

AUTHORS Hell, R.

TITLE Direct Submission

JOURNAL Submitted (14-FEB-2000) R. Hell, Ruhr-Universitaet Bochum, Lehrstuhl fuer Pflanzenphysiologie, Universitaetstr. 150, 44801 Bochum, FRG

COMMENT On Feb 16, 2000 this sequence version replaced gi:602329.

FEATURES

source

1. .1253

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/cultivar="Columbia"

/db\_xref="taxon:3702"

/clone="At.OAS.5-8"

/clone\_1b="lambda ZAP"

/dev\_stage="leaf"

/dev\_1030

/EC\_number="4.2.99.8"

/codon\_start=2

/product="O-acetylserine (thiol) lyase"

/protein\_id="CAA56593.2"

/db\_xref="GI:6983574"

/db\_xref="GOA:P47998"

/db\_xref="SWISS-PROT:P47998"

/translation="MASRIAKDVTBLIGNTPLVYLNNAVAGCVRAKLEMEPCSS VKDRIGFSMISDAEKKGLIKPGESVLIPTSGNTGVLAFTAAAKYKLIITMPASMS TERRILLAFGEVLVITDPAKMGKAIKAEIILAKTPNGYMLQFENPANKIHYET TGEPIWKGTEGIDGFSVSGITGTTGAGKYLKRONANVLYGVEPVESAILSGKPE GPHKIGIGAGRIPSVLNVDLIDVVOVSSDESIDMAROLAKEGULVGISSGAAAAA AIKLAQRPENAGKLVFAIFPSFGERLYSTVLFDATRKAEAMTEEA"

ORIGIN polyA\_signal

1153. .1159

Alignment Scores:

Pred. No.:	Score:	Length:
8.03e-99	1384.00	1253
Percent Similarity:	92.79%	Matches: 266
Best Local Similarity:	83.39%	Conservative: 30
Query Match:	85.27%	Mismatches: 23
DB:	8	Indels: 0
		Gaps: 0

US-09-931-457a-31 (1-325) x ATGACTLY (1-1253)

QY 6 SerGlyTLeAlaLysAspValThrGluLeuIleGlyLysThrProLeuValTyrLeuAsn 25

Db 68 TCGAGATTGCTAAAGATGTGACTGATTGATGGGAACAACCTCCATGGTGTATTGGAAC 127

QY 26 LysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluProCys 45

Db 128 AATGTGCTGAAGATGTGCTGCTGCTGCTGAAGCTTGAAGATGGAACCGTGC 187

QY 46 SerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGlyLysLeu 65

Db 188 TCTAGCGCTCAAGACAGATGCTTGTAGCATGATTCTGATGACAGAAAGAGGTCCTT 247

QY 66 IleThrProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIleGlyLeu 85

Db 248 ATCAACACGAGAGAGATGCTGATTGAGCCCAAGCTGGAACAACCTGAGTTGGGTTA 307

QY 86 AlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProAlaSerMetSer 105

Db 308 GCATTACGCGCAGCTGCTCAAGCTTATTAATTAATGATGAGT 367

QY 106 LeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspProAla 125

Db 368 ACTGAGAGAGAATCATCTCTTACTTTTGAGAGTTGAGTTGTTAACTGACCCAGCT 427

QY 126 LysGlyMetLysGlyAlaValAlaGlnLysAlaGluGluIleLeuAlaLysThrProAsnAla 145

Db 428 AAGGCGATGAAAGAGCTATTCGCAAGCGGAGAGATTTGGCGAAAAACCAATGCT 487

QY 146 TyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThrGly 165

Db 488 TACATGCTTCAGCAGTTTGAGAACCTCCCAACCTTAAGATCCACTATGAGACTACGGGA 547

QY 166 ProGluIleThrPlyGlySerAspGlyLysIleAspAlaPheValSerGlyIleGlyThr 185

Db 548 CCTGAGATATGAGAAAGCACTGCTGCGCAAAATCGATGCTTTGTTCTGGAGTTGTTACT 607

QY 186 GlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGluAsnProAsnIleLysLeu 205

Db 608 GGTGTACCATTAACAGGTGCTGGAGACTATCTTAAGAACAGAACCCAAAGCTCAAGCTG 667

QY 206 IleGlyValGluProValGluSerProValLeuSerGlyLysProGlyProHisLys 225

Db 668 TATGAGTGAGCCAGCTGTAAGTGTATTTCTATTCGGTGGGAGCCAGCTCTCAAGAG 727

QY 226 IleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeuLeuAspGlu 245

Db 728 ATTCAAGGATAGAGAGCTGTTTATTAACCAAGTGTATGATGATGATCTTATTGACGAA 787

QY 246 ValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeuLeuAlaLeuLysGlu 265

Db 788 GTTGTCAAGTTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 847

QY 266 GlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAlaAlaPheGlnIleAlaLys 285

Db 848 GGTCTTCTTGGGAGATATCATCCGCTGACAGCACTGTCAGCAATTAACCTTGACAG 907

QY 286 ArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSerPheGlyLysArgTyr 305

Db 908 AGGCCAGAAACGCTGGAGCTATTGTTGGCGATATCCCGAGTTTGGGAGAGGTTAT 967

QY 306 LeuSerSerValLeuPheGluSerValArgArgGluAlaGluSerMetThrPheGlu 324

Db 968 CTATCAACGCTACTTTTGCATGCGACAGAAAGAGCGGAGGATGACCTTGAG 1024

RESULT 12

AK071279 1337 bp mRNA linear PLN 24-JUL-2003

LOCUS Oryza sativa (japonica cultivar-group) cDNA clone:J023084C02, full insert sequence.

DEFINITION AK071279.1 GI:32981302

ACCESSION AK071279

VERSION AK071279.1

KEYWORDS FLI\_CDNA; CAP trapper.

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)

Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1

AUTHORS The Rice Full-Length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-Length cDNA Project Team; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Oromo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurohaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Nikiura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN; Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,



TITLE  
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.  
Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice  
Science 301 (5631), 376-379 (2003)

JOURNAL  
MEDLINE  
PUBMED  
22752273  
12869764

REFERENCE  
AUTHORS  
2 (bases 1 to 1337)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Nihura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, M., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

TITLE  
JOURNAL  
Direct Submission  
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)  
This clone is one of the 28k full-length cDNA clones from japonica rice.

COMMENT  
URL: http://cdna01.dna.affrc.go.jp/cDNA/  
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

FEATURES  
Source  
1.1337  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultiyar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J023084C02"

FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kuroseki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Nihura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.  
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.  
Location/Qualifiers

ORIGIN  
Alignment Scores:

		Pred. No.:	1.5e-98	length:	1337
		Score:	1381.00	Matches:	264
		Percent Similarity:	89.85%	Conservative:	28
		Best Local Similarity:	81.23%	Mismatches:	33
		Query Match:	85.09%	Indels:	0
		DB:	8	Gaps:	0
		US-09-931-457A-31 (1-325) x AK071279 (1-1337)			
QY	1	MetAlaValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrPro	20		
DB	114	ATGGCGAATCAGGCCAATCGATTGCGAGTATGACTGCACTGATCGCAATACTCC	173		
QY	21	LeuValTyrLeuAsnLysLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu	40		
DB	174	CTGGTGAACCTGAACAAGGTGGATGATGATGAGGCCCAATCGCTGCCAAGCTTGAG	233		
QY	41	LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla	60		
DB	234	ATCATGAGCCCTGCTCCAGTGTCAAGATGATGATTGTTATGATGATTAAGATTAAGATGCA	293		
QY	61	GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn	80		
DB	294	GAAGAGAGGGTCTCATTAATCAGGAAAGAGTGTGATGAGAGCTTACTAGTGAAT	353		
QY	81	ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMet	100		
DB	354	ACTGTAATGGCTTGTCTTCAATGGCTGCTGCCAAGGGTTAAGCTAATCTGCAATG	413		
QY	101	ProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal	120		
DB	414	CCAGCATTCATGAGCATGGAAGAGATCATTTTAAGGCTTCCGTCGAGTGGTC	473		
QY	121	LeuThrAspProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluIleLeuAla	140		
DB	474	CTTACCATCCACTTCTGGGATGAAAGGTGCAATCCAGAGGCAAGATGAGTGGCAGCA	533		
QY	141	LysThrProAsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHis	160		
DB	534	AAGATGCTTAATCTTACATTTCTCCAGCAGTTGAGAACCTGCAAAACCAAGATCCAC	593		
QY	161	TyrGluThrThrGlyProGlnIleThrLysGlySerAspGlyLysIleAspAlaPheVal	180		
DB	594	TATGACCACTGACCTGAGATTGGAAGGCCACTGCTGGGAAAGTGTATATCTTGTTC	653		
QY	181	SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsn	200		
DB	654	TCTGGCATTTGAAACCGGTGTACAGTAACCGGAGCAGGAAGTAACCTCAAGAGCAGAAC	713		
QY	201	ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLys	220		
DB	714	CCTGACATCAAGATCTATGTTGTTGAACCGACAGAAAGTGAATTTATCTGAGAGAGA	773		
QY	221	ProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal	240		
DB	774	CTGGTCCACACAAATCCAGTATAGTGTGTTGTCTGTTGTCTGTTGTTGTTGTTGTTG	833		
QY	241	AsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu	260		
DB	834	AATCTCTGATGATGAGTGTGCAAGTCTCAAGTATGAGGATCAAGATGCGGAGCAG	893		
QY	261	LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla	280		
DB	894	CTAGCGTGAAGAAGAGGCTGCTGTTGGATCTCATCTGCTGCTGCTGCTGCTGCTGCTG	953		
QY	281	PheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSer	300		
DB	954	ATCAGGCTCGCAGCGGCGCCGAGAACAAAGGGAAGCTGTAAGTGTGTTCCCGAGC	1013		
QY	301	PheGlyGluArgTyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSer	320		
DB	1014	TTGGTGAGCGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1073		



QY 321 MetThrpheGluPro 325  
Db 1074 ATGCTGTTCGAGCCA 1088

RESULT 13  
ATCYS3A  
LOCUS ATCYS3A 1332 bp mRNA linear PLN 23-AUG-1995  
DEFINITION A.thaliana mRNA for cysteine synthase.  
X84097  
X84097.1 GI:804949  
KEYWORDS cys-3A gene; cysteine synthase.  
SOURCE Arabidopsis thaliana (chale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE  
1 Barroso, C., Vega, J.M. and Gotor, C.  
A new member of the cytosolic O-acetylserine(thiol) lyase gene  
family in Arabidopsis thaliana  
FEBS Lett. 363 (1-2), 1-5 (1995)  
MEDLINE 95246850  
PUBMED 7729527  
REFERENCE 2 (bases 1 to 1332)  
AUTHORS Gotor, C.  
TITLE Direct Submission  
JOURNAL Submitted (25-JAN-1995) C. Gotor, Inst. de Bioquímica vegetal y  
Fotosint., Dept. de Química, Univ. de Sevilla, A.C.553, B0 41080  
SEVILLA, SPAIN

FEATURES  
source location/Qualifiers  
1..1332  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
1..1332  
/gene="cys-3A"  
99..1067  
/gene="cys-3A"  
/BC\_number="4.2.99.8"  
/codon\_start=1  
/product="cysteine synthase"  
/protein\_id="CA58893.1"  
/db\_xref="GI:804950"  
/db\_xref="GOA:P47998"  
/db\_xref="SWISS-PROT:P47998"  
/translation="MASRIADVTLELIGNTPLYINNVAGCVGRVAAKLEMMPCSS  
VKDRIGFSMISDAERKLPKESVLEPTSGNTGVGLAFTAAKGYKLITMPASMS  
TERRIILAFGVELVLTDPAGKMGKAIKAEELIAKTPNGYMLQGFENPNPKIHET  
TGPEIWKGTGKIDGFPVSGIGTGTITGAGKYLKEONANVKLVGBEPVESAILSGKP  
GPHKIQIGAGFIPSVLNVDLIDEVVOVSSDESIDMARQLAKEGLVGISSGAEEAA  
AIKLAQRPENAGKLFVAIFPSPFGERYLSTVLFDATRKEAEANTPEA"  
234..236  
misc\_binding  
/gene="cys-3A"  
/bound\_moiety="pyridoxal"

Alignment Scores:  
Pred. No.: 2.13e-98 Length: 1332  
Score: 1379.00 Matches: 265  
Percent Similarity: 92.48% Conservative: 30  
Best Local Similarity: 83.07% Mismatches: 24  
Query Match: 84.97% Indels: 0  
DB: 8 Gaps: 0

US-09-931-457A-31 (1-325) X ATCYS3A (1-1332)

QY 6 SerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrProLeuValTyrLeuAsn 25  
Db 105 TCGAAGATTGCTAAAGATGTGCTGAATTGATTGGGAACATTCATGTGTATTGAAAC 164

QY 26 LysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluProCys 45  
Db 165 AATGTTCTGAAGATGTGTGTGTGTGTGTGTGTGTGAAGCTTGAGATGATGAACCGTGC 224

QY 46 SerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluLysGlyLeu 65  
Db 225 TCTAGCGTCAAGACAGGATGTGTTTACATGATTTCTGATGCAGAAAGAGGCTCT 284

QY 66 IleThrProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIleGlyLeu 85  
Db 285 ATCAACCAGAGAGAGTGTCTGATTGACCAACAGTGAACACTGAGCTTGCGTTA 344

QY 86 AlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProAlaSerMetSer 105  
Db 345 GCATTCAAGCGAGCTGCCAAGGCTACAGCTTATTATTAATGCAATGCCAGCTTATGAGT 404

QY 106 LeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspProAla 125  
Db 405 ACTGAGAGAGAAATCATCTCTTAGCTTTTGAGTTGAGTTGAGTTGTTTAATGACCCAGCT 464

QY 126 LysGlyMetLysGlyAlaValAlaGlnLysAlaGluGluIleLeuAlaLysThrProAsnAla 145  
Db 465 AAGGCATGAAGAGAGCTATCGCAAGCGGAGAGATTGTGGCAAAACCAACCAATGCT 524

QY 146 TyrIleLeuGlnIlePheGluAsnProAlaAsnProLysValHisTyrGluThrThrGly 165  
Db 525 TACATGCTTCAGCAGTTTGAGAACCTGCCAACCTTAAGATCCACTATGAGACTACGGA 584

QY 166 ProGluIleThrLysGlySerAspGlyLysIleAspAlaPheValSerGlyIleGlyThr 185  
Db 585 CCTGAGATATGAAAGCACTGCTGCGCAAAATCGATGGGTTTGTTCGGAATGTGACT 644

QY 186 GlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsnProAsnIleLysLeu 205  
Db 645 GGTGATACCATTTACAGTGTCTGGAGAGTATCTTAAGAACAGAACGCAAGCTCAAGCTG 704

QY 206 IleGlyValGluProValGluSerProValLeuSerGlyLysProGlyProHisLys 225  
Db 705 TATGAGTGAAGCCAGTTGAAGTCTATTCTATCCGGTGGAGGAGCCAGTCTCTCAAG 764

QY 226 IleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeuLeuAspGlu 245  
Db 765 ATTCAAGGATAGAGAGCTGTTTATACCAAGTGTATTGAATGTTGATCTTATGACGAA 824

QY 246 ValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeuLeuAlaLeuLysGlu 265  
Db 825 GTTGTTCAGTTTCAAGTATGATCAATCCATTGACATGGCAAGGAGCTTCTTAAGAA 884

QY 266 GlyLeuPheValGlyLysSerSerGlyAlaAlaAlaAlaAlaPheGlnIleAlaLys 285  
Db 885 GGTCTTCTTGTGGGAATATCATCCGGTGCAGAGCTGCTGACCAATTAACTTGACACAG 944

QY 286 ArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSerPheGlyGluArgTyr 305  
Db 945 AGCCAGAAACGCTGGAGAGCTATTGTGGCAATFTCCGAGTTTCGGGAGAGGCTAT 1004

QY 306 LeuSerSerValLeuPheGluSerValArgArgGluAlaGluSerMetThrPheGlu 324  
Db 1005 CTATCAACGCTACTTTTCATGATCGACCAAGAGAAAGCGGAGCAATGACTTTCAG 1061

RESULT 14  
ZMCSOATL 1319 bp mRNA linear PLN 28-JAN-2003  
LOCUS ZMCSOATL 1319 bp mRNA linear PLN 28-JAN-2003  
DEFINITION Z.mays mRNA for cysteine synthase.  
X85803  
X85803.1 GI:758352  
KEYWORDS cysteine synthase; Mcy3p gene; O-acetylserine(thiol) lyase.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
ciade; Panicoideae; Andropogoneae; Zea.

REFERENCE  
1 Brander, K.A., Owttrim, G.W. and Brunold, C.  
Isolation of a cDNA (EMBL X85803) encoding a putative chloroplastic  
isoform of cysteine synthase from maize (PGR95-031)

JOURNAL Plant Physiol. 108, 1748-1748 (1995)  
REFERENCE 2 (bases 1 to 1319)  
AUTHORS Blander, K.A.  
TITLE Direct Submission  
JOURNAL Submitted (23-MAR-1995) K.A. Blander, Institute of Plant  
Physiology, Altenbergstr. 21, CH-3013 Bern, SWITZERLAND

FEATURES  
source  
1.1319  
/organism="Zea mays"  
/mol\_type="mRNA"  
/cultivar="B73 inbred line"  
/db\_xref="taxon:4577"  
/tissue\_type="upper 12 root"  
/dev\_stage="15 weeks"  
1.1319  
/gene="Mcygp"  
73.1050  
/gene="Mcygp"  
/EC\_number="4.2.99.8"  
/note="cysteine synthase"  
/codon\_start=1  
/product="O-acetylserine (thiol) lyase"  
/protein\_id="CA59798.1"  
/db\_xref="GI:758353"  
/db\_xref="GOA:P80608"  
/db\_xref="SWISS-PROT:P80608"  
/translation="MGESAPSIAXDVTLLGNTPLVYLNKVTLDGCVGRRALESMEP  
CSSVKDRIGYSMTDAEEKGLITPGSVLIBTSNTGIGLAFMAAKGYQLTLTPPA  
SSMERRIILKAGAEVLVLTPLGKGAVKRHSIQAKTENSYLQGFENPANKI  
YETGPPIWKATAGKIDGLVSGIGTGTTGTGRYLRQNPVKLYGVEPVSAVLNG  
GKPGPHXIGIGAGTIPGLVDVLDLDTLOVSSDAIETAKALAKEGLVGISGA  
AAAVRLAKRPENAGKLFVVFVPSFGERYLSVLFOSIKKEASMTVEP"  
73.180  
/gene="Mcygp"  
181.1047  
/gene="Mcygp"  
/product="unnamed"

sig\_peptide  
mat\_peptide

ORIGIN

Alignment Scores:  
Pred. No.: 2.52e-98 Length: 1319  
Score: 1378.00 Matches: 268  
Percent Similarity: 91.19% Conservative: 22  
Best Local Similarity: 84.28% Mismatches: 28  
Query Match: 84.90% Indels: 0  
DB: 8 Gaps: 0

US-09-931-457A-31 (1-325) x ZWCSQATL (1-1319)

QY 8 ILEAlAlAspValThrgluLeuileGlyLysThrProLeuValTyrLeuAsnLysLeu 27  
|||  
DB 94 ATGCCCAAGACGTCAACGAGTTGATCGGGAACACGCCGCTGCTACTCAACAAGTG 153  
|||  
QY 28 AIAspGlyCySvalAlaArgValAlaAlaLysLeuGluLeuMetGluProCySserSer 47  
|||  
DB 154 ACCGATGGGTGCGTCCGCCGACGTCCGCCAAGCTCGAGTCCATGAGAGCCCTGCCAGC 213  
|||  
QY 48 VallysAspArgileGlyTyrSerMetileAlaAspAlaGluGluLysGlyLeuileThr 67  
|||  
DB 214 GTCAAGATAGGATGTGGCTACAGCATGATCAACGACGACAGAGAGAAGGGCTGATTACT 273  
|||  
QY 68 ProGlyLysSerValLeuileGluProThrSerGlyAsnThrGlyileGlyLeuAlaPhe 87  
|||  
DB 274 CCAGGCGTGAAGTCTTCTGATTGAACCAACTAGCCGCAACAGGCATTGAGCTGGCTTT 333  
|||  
QY 88 MetaAlaAlaArgGlyTyrLysLeuileileThrMetProAlaSerMetSerLeuGlu 107  
|||  
DB 334 ATGGCTGCTGCCAAGGCTTACAACTTACACTCACAATGCTGCTCCATGAGCATGAG 393  
|||  
QY 108 ArgArgileileLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspProAlaLysGly 127  
|||  
DB 394 AGGAGGATCATATTGAAGGCTTTTGTGCTGCTGAACCTGTCTCTTACTGACCCACTCTTGGGA 453  
|||

**gene**

```
128 MetLysGlyAlaValGlnIleLysAlaGlutLeuAlaHisThrProAsnAlaTyrTrl 147  
|||:::||:  
454 ATGAAGACCTGTCAAGAAACCGAAGATACAAGCACCAGAACCCTACTGGTCATT 513  
  
148 LeuGlnGlnPheGluAsnProlAsnProLysValHistYrghuThrhrglyProglu 167  
||||  
514 CTCCAACCATTTGAAAATCCAAGCTTAACCCAAAGATTCAATAAGACTACTGGGCCCTGAA 573  
  
168 IleTrypLysSeraspSilylsileaspaIapheValSerGlyIleGlyThrglygly 187  
||||::|  
574 ATTGGAAAGCTACAGCAAAGAAAAATTGATGCCTTGATCTGSTATCGGACAGAGGT 633  
  
188 ThrIlethrGlyalaGlylysTyrlauylsglnginsnpRoasnileylsleulegity 207  
|||||::|  
634 ACCATCACCGGTACTGGCGCATACCTCAGAGAGCAAATCTAATGTCAAGCTTATGGT 693  
  
208 ValGluProvalGluSerProvalLeuSerGlyglysproglyProhislysilegin 227  
|||||  
694 GTGAGCCAGTGGAGAGTGTCTGTTTGAATGGTGAAAAACCTGGACACACAAGATTCAA 753  
  
228 GlyIleglyalaglypherleproglyVallaugiuValaslneuleuaspgluValval 247  
|||||  
754 GGAATGGAGCTGGTTTTATCCCTCGAGTCTTGATGTGATCTCCTGATGAATCTTA 813  
  
248 GlulleSerSeraspGluAlalegluthralalsleuleuAlaleulysgluglyleu 267  
|||::|  
814 CAGTTTCAAGTAGTAGAGCTATCGAGACTGCCAAGGCTCTGCTCTGAAGAAGSGTTG 873  
  
268 PheValGlylleSerSerGlyAlaalaaialaalalapheGlnIleAlalyArgPro 287  
|||||  
874 TTGGTGGATCTCTTCTGTGTGAGCTGCAGCTGGCGGAGTGAAGCTGTGTAAGAGCCA 933  
  
288 GluAsnaIeagllyslenuileValAlavalpheProserPheGlygluArgtyrleuSer 307  
|||||  
934 GA AACGCCGGAAGCTATTGTTGTTGTCGTGTCCAGCTTGGCGGAGCGTCACTCTCA 993  
  
308 SerValleuphegluservalargglualaglusermethrphregluPro 325  
|||||  
994 TCGGTCTGTTCCAGTCCATCAAGAAGAGCGGACGATGTGTGAGAGCCC 1047  
  
Db  
  
Qy  
  
RESULT 15  
AF073697 1363 bp mRNA linear PLN 08-APR-1999  
LOCUS AF073697  
DEFINITION Oryza sativa cysteine synthase (rcs3) mRNA, complete cds.  
ACCESSION AF073697  
VERSION AF073697.1 GI:4574138  
KEYWORDS  
SOURCE  
ORGANISM  
.Oryza sativa {japonica cultivar-group}  
Oryza sativa {japonica cultivar-group}  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoidae; Oryzaceae; Oryza.  
REFERENCE  
AUTHORS Nakamura,T., Yamaguchi,Y. and Sano,H.  
TITLE Four rice genes encoding cysteine synthase: isolation and differential responses to sulfur, nitrogen and light  
JOURNAL Gene 229 (1-2), 155-161 (1999)  
MEDLINE 99196994  
PubMed 10095115  
REFERENCE  
AUTHORS Nakamura,T.  
TITLE Direct Submission  
JOURNAL Submitted (23-JUN-1998) Research and Education Center for Genetic Information, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0101, Japan  
FEATURES  
source location/Qualifiers  
1..1363  
    /organism="Oryza sativa (japonica cultivar-group)"  
    /mol_type="mRNA"  
    /cultivar="Nipponbare"  
    /db_xref="taxon:39947"  
1..1363  
    /gene="rcs3"
```

CDS  
139..1116  
/gene="rcs3"  
/EC\_number="4.2.99.8"  
/note="O-acetylserine (thiol)-lyase"  
/codon\_start=1  
/product="cysteine synthase"  
/protein\_id="AAD23909.1"  
/db\_xref="GI:4574139"  
/translation="MAESGQSIASDVTAIGNTPLVYLNKVVDCEAOIAKLRIEMPCSSVKDRIGYSMTDAEKGITPGKSVLEPTSGNTGIGLAFMAAKGYKLITMPASMSMERRIILKAFGAELVLTPLGMKGAIOKADBLAAMPNSTILOFENPANKIHYETTGPEIWKATAGKVDILVSGIGTGTGKYLKEONPEIKIYGVETPESAIISGGRPGPHKIQIGAGFVPGVLDVNLDEVVQVSSDBAISMAKQIALKEGLLVGISGAAAVAIRVAQRPENRKGLVVVFPSPFGERYLSVLFBSIKREANMVEEP"

ORIGIN

Alignment Scores:  
pred. No.: 3.14e-98 length: 1363  
Score: 1377.00 matches: 263  
Percent Similarity: 89.54% conservative: 28  
Best Local Similarity: 80.92% mismatches: 34  
Query Match: 84.84% indels: 0  
DB: 8 gaps: 0

US-09-931-457a-31 (1-325) x AF073697 (1-1363)

QY 1 MetAlaValGluArgSerGlyLeuAlaLysAspValThrGluLeuIleGlyLeuThrPro 20  
DB 139 ATGCCGGAATCAGGCCAATCGATTGCGAGTGTGACTGACATCGCACTGCGCAATACTCCC 198  
QY 21 LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40  
DB 199 CTGGTGTACTGTAACAAGGTGGTGGATGATGTGAGGCCCAATCGCTGCCAAGCTTGAG 258  
QY 41 LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60  
DB 259 ATCATGAGCCCTGCTCCAGTGTCAAGATAGGATTGGTTATAGCATGATTACAGATGCA 318  
QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80  
DB 319 GAAGAGAAAGGCTCTCATTTACTCCAGGAAAGAGTGTTTGATTGAGCCCTACTAGTGAAT 378  
QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMet 100  
DB 379 ACTGGTAFTGGCTTGTCTTTCAATGGCTGCTGCCAAGGCTTATAAGCTTAATCTGACAA 438  
QY 101 ProAlaSerMetSerLeuGluArgGlyIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120  
DB 439 CCAGCATCCATGACATGGAAGAGGATCATTTCTTAAGGCTTGGGTGCCGAGTTGCTC 498  
QY 121 LeuThrAspProAlaLysGlyMetLysGlyAlaValGluLysAlaGluGluLeuAla 140  
DB 499 CTTACCGATCCACTTCTGGGCAATGAAGGTGCATTCAGAGGCGAGATGAGCTGGCAGCA 558  
QY 141 LysThrProAsnAlaTyrIleLeuGluGlnPheGluAsnProAlaAsnProLysValHis 160  
DB 559 AAGATGCTTAACCTTACATTTCTCCAGAGTTGAGAACCTTGCAAAACCCAAAGATCCAC 618  
QY 161 TyrGluThrThrGlyProGluIleThrLysGlySerAspGlyLysIleAspAlaPheVal 180  
DB 619 TATGAGACAACTGACCTGAGATTGGAAGGCCACTGCTGGGAAAGTTGATATCTTGTTC 678  
QY 181 SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsn 200  
DB 679 TCTGGCATTTGGAACCGGTGTGACGTAAACCGGAACAGAAAGTACCTCAAGGAGCAGAAC 738  
QY 201 ProGluIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLys 220  
DB 739 CCTGAGATCAAGATCTATGGTGTGAACCGACAGAAAGTGCATTTATCTGAGGAGAGA 798  
QY 221 ProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240

QY 241 AsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260  
DB 859 AATCTCTGATGAAGTGTCTCCAGTCTCAAGTGAAGCCATGACATGGCGAAGCAG 918  
QY 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla 280  
DB 919 CTAGCGTTGAAGAGAGGCTGCTGGTTGGGATCTCATCTGCTGCTGCTGCTGCTGCC 978  
QY 281 PheGluIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaAlaPheProSer 300  
DB 979 ATCAGGTCGCGAGCGCGCCGAGAACCAAGGAACTCGTAGTTGTGTCTTCCCGAGC 1038  
QY 301 PheGlyLysArgTyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSer 320  
DB 1039 TTCGGTAGCGCTAAGTGTCTGTCTCTTCTTCAATCATGATCAAGAGGAGCGCAAAAC 1098  
QY 321 MetThrPheGluPro 325  
DB 1099 ATGGTGTTCAGCCA 1113

Search completed: May 21, 2004, 23:43:15  
Job time : 4706 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 21, 2004, 21:58:57 ; Search time 3100 Seconds  
(without alignments)  
3130.712 Million cell updates/sec

Title: US-09-931-457A-31  
Perfect score: 1623  
Sequence: 1 MAVERSGIAKVTTELIGKTP.....ISSVLFESVRRRAESMTPEP 325

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DBV=x1h  
-Q=/cgn2\_1/USPTO\_spool/US09931457/runat\_18052004\_121729\_20307/app\_query.fasta\_1.519  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09931457 @CGN 1.1 1906 @runat\_18052004\_121729\_20307 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pin:\*  
20: em\_gss\_vtc:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mub:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vtl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1417	87.3	994	14	CK260011 EST706089
2	1392	85.8	929	14	CF513886 Cabud0007
3	1389	85.6	1483	11	AY103651 Zea mays
4	1372	84.5	967	14	CK271210 EST717288
5	1367	84.2	910	14	CB974516 CAB30005
6	1364	84.0	937	14	CK266154 EST712232
7	1313	80.9	881	14	CK259489 EST705567
8	1311	80.8	907	14	CK260012 EST706090
9	1311	80.8	920	14	CK259857 EST705935
10	1310	80.7	873	14	CK259488 EST705566
11	1306	80.5	917	14	CK259858 EST705936
12	1291	79.5	935	14	CK279506 EST725584
13	1285	79.2	934	29	CG848066 ZMMBB031
14	1269	78.2	1258	11	AY112530 Zea mays
15	1260	77.6	886	14	CK265667 EST711745
16	1250	77.0	802	14	CA801406 Bau05C03
17	1221	75.2	802	14	CB974696 CAB30005
18	1214	74.8	845	14	CF200608 RR890915N
19	1214	74.8	1155	14	CK162330 FGAS01492
20	1206	74.3	792	14	CF512887 CABud0005
21	1197	73.8	965	14	CK274117 EST720195
22	1194.5	73.6	803	12	BG648907 EST510526
23	1192	73.4	1111	14	CK163592 FGAS01622
24	1186	73.1	724	10	BF003458 EST431956
25	1186	73.1	813	14	CF205785 RR890915I
26	1174	72.3	791	12	BG595822 EST494500
27	1172	72.2	888	14	CK266746 EST712824
28	1156	71.2	856	14	CF652605 65-L02057
29	1155	71.2	914	14	CD439033 BL01N0520
30	1149	70.8	730	13	BU027058 OHG18P11
31	1138	70.1	785	12	BG587892 EST489667
32	1117	68.8	866	14	CD435997 BL01N0369
33	1110	68.4	673	12	BM815595 EST593689
34	1107	68.2	734	12	BG648870 BST510489
35	1104	68.0	726	12	BG128810 EST474456
36	1088	67.0	742	14	CD813753 BN15.0200
37	1084	66.8	655	12	B1309270 BST530680
38	1081	66.6	740	13	BQ914877 QHB12F18
39	1075	66.2	804	14	CB658211 OSJNEC14B
40	1073	66.1	694	14	CK273767 EST719845
41	1071	66.0	686	12	B1264323 NF118G04P
42	1070	65.9	678	10	AW311624 sg43b07.Y
43	1068	65.8	800	14	CB651906 OSJNEC01B
44	1066	65.7	806	14	CB654896 OSJNEC07M
45	1064	65.6	768	14	CD824490 BN25.053L

ALIGNMENTS

RESULT 1  
LOCUS CK260011 994 bp mRNA EST 12-DEC-2003  
DEFINITION EST706089 potato abiotic stress cDNA library Solanum tuberosum cDNA  
clone POAB228 5' end, mRNA sequence.  
ACCESSION CK260011  
VERSION CK260011.1 GI:39816989  
KEYWORDS EST.  
SOURCE Solanum tuberosum (potato)  
ORGANISM Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum.



REFERENCE 1 (bases 1 to 994)  
AUTHORS Buell, C.R., Hart, A., Zismann, V., Karanymcheva, S.A. and Baker, B.  
TITLE Generation of ESTs from abiotic stressed potato tissue  
JOURNAL Unpublished (2003)  
COMMENT Other ESTs: EST706088 EST706090 EST706091  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from TIGR via potato@tigr.org  
Seq primer: ATT TAG GCG ACA CTA TAG.  
Location/Qualifiers

FEATURES  
source 1. .994  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="POAB228"  
/tissue\_type="abiotic stress treated leaf and root tissue"  
/lab\_host="DHL08-Tona"  
/clone\_lib="potato abiotic stress cDNA library"  
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: Solanum tuberosum var. Kennebec plants were grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:  
Pred. No.: 5.44e-143 Length: 994  
Score: 1417.00 Matches: 279  
Percent Similarity: 94.10% Conservative: 24  
Best Local Similarity: 86.65% Mismatches: 19  
Query Match: 87.31% Indels: 1  
DB: 14 Gaps: 0

US-09-931-457a-31 (1-325) x CR260011 (1-994)

QY 1 MetAlaValGluArgSerGlyIleAlaIysAspValThrGluLeuIleGlyIysThrPro 20  
DB 28 ATGGCGGGGAAAGATTGGATTGCCAAGATGTAATGATTGTTAACAACCTCT 87  
QY 21 LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaIaIysLeuGlu 40  
DB 88 TTGGTATACCTGAATAATGTTGGATGGGTGTGTGCACTGTTGCTGCCAAGCTGAA 147  
QY 41 LeuMetGluProCysSerSerValIysAspArgIleGlyTyrSerMetIleAlaAspAla 60  
DB 148 AGCATGAGCCATGCTCTGAGTTAAGATAGATCGGTTAATGATGATTACAGATGCT 207  
QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80  
DB 208 GAGGAGAAGGCTCTCATCAACCCGCGAGAGTGTCTCATGAACTACGAGTGAAGAAC 267  
QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaIaArgGlyTyrLysLeuIleIleThrMet 100

DB 268 ACTGCTGTAGGATTGGCATTCATGCTGCTGCTAAAGGCTACAACTCATCATTCAGATG 327  
QY 101 ProAlaSerMetSerLeuGluArgGlyIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120  
DB 328 CCTCTTCAATGAGTCTTGAGAGAGAATATTCTGCGCTTTCGGTCTGAGTTGGTG 387  
QY 121 LeuThrAspProAlaIysGlyMetLysGlyAlaValGluIysAlaGluIleLeuAla 140  
DB 388 CTTACCGATCCAGCAAAAGGATGAAGGCTCTATTTCAAAGGCTGAAGATTAAGGCC 447  
QY 141 LysThrProAsnAlaTyrIleLeuGluIlePheGluAsnProAlaAsnProLysValHis 160  
DB 448 AAAACACCCCACTCCTATATTCTTCAGCAATTGAAAACCTCTTACCCAAAGATACAC 507  
QY 161 TyrGluThrThrGlyProGluIleTrpLysGlySerAspGlyLysIleAspAlaPheVal 180  
DB 508 TATGAGACCACTGCTCTGAGATCTGAAAGGCTCAATGGGAAAGTGATGCTTAGTC 567  
QY 181 SerGlyIleGlyThrGlyLysThrIleThrGlyAlaGlyLysTyrLeuLysGluGluAsn 200  
DB 568 TCTGGAATTGGAACAGAGGACAGTAACTGTTCAAGGCAAGTATTGAGAGAGCAGAAC 627  
QY 201 ProAsnIleLysIleuIleGlyValGluProValGluSerProValLeuSerGlyLys 220  
DB 628 CCCAACGTTAAGCTGTATAGCGCGTGAACCAAGTGAAGTGTATCTTCTGTTGGAAG 687  
QY 221 ProGlyProHisLysIleGluGlyAlaGlyPheIleProGlyValLeuGluVal 240  
DB 688 CCGTGTCCACATTAAGATTCAAGGGAATGCTGCTGTTCAATCTGCTGTTTGAAGTT 747  
QY 241 AsnLeuLeuAspGluValValGluIleSerSerAspGluAlaIleGluThrAlaLysLeu 260  
DB 748 AACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 807  
QY 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla 280  
DB 808 CTGCATTAAGAAAGGATGCTAGTGAATTTCACTGCTGCTGCTGCTGCTGCTGCTGCT 867  
QY 281 PheGluIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSer 300  
DB 868 ATTAAGTGTCTAAGCGCCCTGAAGATCTGGAAGCTCATTTGTTAT-TTCCACAGC 926  
QY 301 PheGlyLysArgTyrLeuSerSerValLeuPheGluSerValArgGluAlaGluSer 320  
DB 927 TTGGAGAGGATATCTTCTCTGTAATCTTCTGAACTGTCAAGACGAGAGACAGAGAAC 986  
QY 321 MetThr 322  
DB 987 ATGACT 992

RESULT 2  
CF513886 929 bp mRNA linear EST 09-SEP-2003  
LOCUS Cabud0007 IIIF\_H10 Vitis vinifera cv. cabernet sauvignon (Clone 8)  
DEFINITION Bud - CABUD Vitis vinifera cDNA clone Cabud0007\_IIIF\_H10 5', mRNA  
sequence.

ACCESSION CF513886  
VERSION CF513886.1 GI:34545654  
KEYWORDS EST.  
SOURCE Vitis vinifera  
ORGANISM Vitis vinifera  
Vitis vinifera  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; Vitaceae; Vitis.  
REFERENCE 1 (bases 1 to 929)  
Goes da Silva, F., Iandolo, A., Lim, H., Baek, J., Jones, K. and  
Cook, D.  
Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'  
berries at various developmental stages  
Unpublished (2003)  
COMMENT JOURNAL  
CONTACT: Douglas Cook, PhD  
CABES Genome Facility

UC Davis, Plant Pathology  
One Shields Ave, Davis, CA 95616, USA  
Tel: 530 754 6561  
Fax: 530 754 6617  
Email: drcook@ucdavis.edu  
Seq primer: ACGGTACCGACATATGCC.  
Location/Qualifiers

## FEATURES

source

1. .929  
/organism="Vitis vinifera"  
/mol\_type="mRNA"  
/cultiivar="Cabernet Sauvignon (Clone 8)"  
/db\_xref="taxon:29760"  
/clone="Cabud0007\_11IF\_H10"  
/sex="Hermaphrodite"  
/dev\_stage="Pre-bloom (10-11 days before bloom)"  
/lab\_host="VH5alpha"  
/clone\_lib="Vitis vinifera cv. cabernet sauvignon (Clone 8) Bud - CABUD"  
/note="Organ: Bud; Vector: pDNR; Site\_1: S11; Site\_2: S11; CABUD is a cDNA library of Vitis vinifera cv. Cabernet Sauvignon. Clone 8 dissected buds. Samples were collected May 13, 2002 from pre-bloom plants (10-11 days before bloom), pre-veraison. Sampled vines were located at the University of California, Davis, Experimental Vineyard. cDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows:  
5'-AAGCAGTGTATCAACGACGAGTGGCCATTACGCCGG-3' and  
5'-ATTCTAGAGCGCGGCGCGGCGGACATG-dT(30)NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

## ORIGIN

## Alignment Scores:

Pred. No.: 2.53e-140 Length: 929  
Score: 1392.00 Matches: 269  
Percent Similarity: 95.74% Conservative: 23  
Best Local Similarity: 88.20% Mismatches: 13  
Query Match: 85.77% Indels: 0  
DB: 14 Gaps: 0

US-09-931-457a-31 (1-325) x CFS13886 (1-929)

QY 21 LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40  
|||||  
Db 1 TTAGTATATCTGAACATGCTGTAGATGCTGTGTACTCTGATTGCTGCAAGCTTGA 60  
QY 41 LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60  
:::|||||  
Db 61 ATGATGAGCCTTGTCTGTAGTGTCAAGACAGATCGGCTATAGTATGATGATCAAGATGCA 120  
QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValIleGluProThrSerGlyAsn 80  
|||||  
Db 121 GAAGAGAAGGCGCTCATTAAGCCAGGAGAGAGTGTCTCATTAAGCTTACCAAGTGTAT 180  
QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysIleuIleThrMet 100  
|||||  
Db 181 ACTGGCATAGGCTTGTGCTTATGCGCGCAGCTAAGGTTATAGCTCATCATTAACAATG 240  
QY 101 ProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120  
|||||  
Db 241 CCTGCTTCGATGAGTCTTGAAGAAGATGCTTCCGAGCTTTGGCGCTGAAATGTT 300  
QY 121 LeuThrAspProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluGluIleLeuAla 140  
|||||  
Db 301 CTCAACAGATCCAGCCCAAGGCGCATGAAGGGGCTGTTCAAGAAAGCTGAAGAGATCGCA 360  
QY 141 LysThrProAsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHis 160  
|||||  
Db 361 AAGACACCTAATTCCTATATCTTCAAGCAATTTGAACCCCTGCCAATCCGAAGATTCAT 420  
QY 161 TyrGluThrThrGlyProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheVal 180  
|||||

Db 421 TATGAACCACTGGACCCGAGATCTGGAGAGCAGCACTGTGGAAAGTTGATCTTTGTC 480  
QY 181 SerGlyIleGlyThrGlyGlyTyrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsn 200  
|||||  
Db 481 TCTGGATAGGACTGGAGGTACCATATACAGGTGACGGAAGTTCTCAAGACCAAAAT 540  
QY 201 ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLys 220  
:::|||||  
Db 541 TCTGACATTAGCTGTATGTGTAGAACTGTGAAAGTGCAGTACTGTCTGAGAGAAA 600  
QY 221 ProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240  
|||||  
Db 601 CTTGCGCCGCTAAGATCCAGGAATTCGCGCTGTTTCAATCCCTGAGATTAGATGTC 660  
QY 241 AsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260  
|||||  
Db 661 AATTGCTTGTATGAACTTGTCAAGTTTCAAGTGAAGAGCTGTGAAGCTT 720  
QY 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla 280  
|||||  
Db 721 CTGCAATTGAAGAAGGTTTGTCTGCTGTGGATATATCTGTGTGTGCTGCACTGCTGCA 780  
QY 281 PheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSer 300  
:::|||||  
Db 781 ATAAAGTTAGCAAGAGCGCAGAAATCTGGGAAACTCATTTGTTGTTTCCCAAGC 840  
QY 301 PheGlyGluArgTyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSer 320  
|||||  
Db 841 TTGGAGAGCTTATCTGTCTGCTGTGTGTGCTGCTGTGAGCTGTGAGCGGAGGCAAGAAAT 900  
QY 321 MetThrPheGluPro 325  
|||  
Db 901 ATGCTGTTGAACCC 915

RESULT 3  
AY103651 1483 bp mRNA linear HTC 16-OCT-2002  
LOCUS  
DEFINITION  
Zea mays PC0145346 mRNA sequence.  
AY103651  
VERSION  
AY103651.1 GI:21206729  
KEYWORDS  
SOURCE  
ORGANISM  
Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE  
AUTHORS  
1 (bases 1 to 1483)  
Bainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,  
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.  
Maize Mapping Project/Dupont Consensus Sequences for Design of  
Overgo Probes  
Unpublished (2002)  
2 (bases 1 to 1483)  
Coe, E.H.

JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Direct Submission  
Submitted (25-APR-2002) Maize Mapping Project, University of  
Missouri, Columbia, MO 65211, USA  
If you are interested in getting corresponding physical clones,  
these are publicly available from ZmDB, www.zmdb.iastate.edu; TIGR,  
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the  
maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
Schnable, Iowa State, then clones may be requested from ZmDB:  
www.zmdb.iastate.edu.

FEATURES  
source  
1. .1483  
location/Qualifiers

/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="MaizeDB:638596"  
/db\_xref="taxon:4577"  
/clone\_lib="Maize Mapping Project/Dupont Consensus  
Library"  
/note="this sequence is part of a project of EST

assemblies resulting from the application of public  
contigs to seed Dupont contigs; this resource was  
assembled by Dupont as part of a collaboration for the  
overgo addressing of BACS in conjunction with the Maize  
Mapping Project"

ORIGIN

Alignment Scores:

Pred. No.:	1,16e-139	Length:	1483
Score:	1389.00	Matches:	270
Percent Similarity:	91.82%	Conservative:	22
Best Local Similarity:	84.91%	Mismatches:	26
Query Match:	85.58%	Indels:	0
DB:	11	Gaps:	0

US-09-931-457a-31 (1-325) \* AY103651 (1-1483)

QY 8 ILEAlaLysAspValThrGluLeuILEGlySThrProLeuValTyrLeuAsnLysLeu 27  
DB 242 ATGCCAAGGACGTACCGAGTTGATCGGACACGCCGCTGTACTCAACAAGTG 301  
QY 28 AlaAspGlyCysValAlaArgValAlaAlaLysLeuGluMetGluProCysSerSer 47  
DB 302 ACCGATGGGTGCGTCGCGCGCGTCGCCCAAGCTCGAGTCCATGAGCCCTGCTCCAGC 361  
QY 48 ValLysAspArgILEGlyTyrSerMetILEAlaAspAlaGluGluLysGlyLeuILEThr 67  
DB 362 GTCAAGGATAGATTGCTACACATGATCAGGACGCGAGAGAGAAAGGCTGATTACT 421  
QY 68 ProGlyLysSerValLeuILEGluProThrSerGlyAsnThrGlyILEGlyLeuAlaPhe 87  
DB 422 CCAGCGGTAGTGTGATTGAAACCACTAGCGGCAACAGGCATTGAGCTGCTTT 481  
QY 88 MetAlaAlaArgGlyTyrLysLeuILEThrMetProAlaSerMetSerLeuGlu 107  
DB 482 ATGGCTGCTGCCAAGGCTTACAACTTACATCAATGCTGCTCCATGAGCATGGAG 541  
QY 108 ArgArgILELeuLeuAlaPheGlyAlaGluLeuValLeuThrAspProAlaLysGly 127  
DB 542 AGGAGATCATATTGAAGCTTTTGTGCTGAGCTTGTCTTACTGAGCCCACTCTTGGGA 601  
QY 128 MetLysGlyAlaValGlnLysAlaGluLysLeuAlaLysThrProAsnAlaTyrILE 147  
DB 602 AFGAAGAGAGCTGTCAAGAAAGCGAGAGATACAAAGAAAGACACCACTGTCATCATT 661  
QY 148 LeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThrGlyProGlu 167  
DB 662 CTTCAACAATTGAAATCCAGCTAACCCAAAGATTCATGAGACTACTGGGCTTGA 721  
QY 168 ILETrpLysGlySerAspGlyLysILEAspAlaPheValSerGlyILEGlyThrGlyGly 187  
DB 722 ATCTGGAAGACTACAGCAAGAAATGATGCGCTTGTATCTGTATCGGACAGAGGT 781  
QY 188 ThrILEThrGlyAlaGlyTyrTyrLeuLysGluGlnAsnProAsnILELysLeuILEGly 207  
DB 782 ACCATCAACCGGTACTGGGCGATACCTCAGAGAGCAAAATCTTAATGTCAAGCTCTATGCT 841  
QY 208 ValGluProValGluSerProValLeuSerGlyLysProGlyProHisLysILEGln 227  
DB 842 GTGAGCCAGTGTGAGAGTGTCTTTGATGTGAAAACTCGAACACACAGATTCAA 901  
QY 228 GlyILEGlyAlaGlyPheILEProGlyValLeuGluValAsnLeuLeuAspGluValVal 247  
DB 902 GGAATTGAGCTGTTTATCTCTGAGTCTTGATGATCTCTCTGATGAACCTCTA 961  
QY 248 GlnILESerSerAspGluAlaILEGluThrAlaLysLeuLeuAlaLeuLysGluGlyLeu 267  
DB 962 CAGGTTTCAAGTATGAGCTATTCAGAGACTGCCAAGGCTCTGCTGAAAGAGGCTTG 1021  
QY 268 PheValGlyLysSerSerGlyAlaAlaAlaAlaAlaPheGlnILEAlaLysArgPro 287  
DB 1022 TTGGTTGGAATCTCTTCTGTGTGACGTGACGCTGCGGCACTTAAGCTTGAAGGCCCA 1081

QY 288 GluAsnAlaGlyLysLeuILEValAlaValPheProSerPheGlyGluArgTyrLeuSer 307  
DB 1082 GAAAGCCCGGAAAGCTATTGTGTGCTGTGCTCCGAGCTTCGGCAGCGCTACTCTCA 1141  
QY 308 SerValLeuPheGluSerValArgGluAlaGluSerMetThrPheGluPro 325  
DB 1142 TCGGTGCTGTTCCAGTCCATCAAGAGAGCGGAGCATGTGTGAGGCC 1195

RESULT 4

CK271210 967 bp mRNA linear EST 12-DEC-2003  
DEFINITION EST717288 potato abiotic stress cDNA library Solanum tuberosum cDNA  
clone POACX82 5' end, mRNA sequence.

CK271210.1 GI:39828188  
EST.  
Solanum tuberosum (potato)

ORGANISM Solanum tuberosum  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum.

REFERENCE 1 (bases 1 to 967)  
AUTHORS Buell,C.R., Hart,A., Zismann,V., Karamychova,S.A. and Baker,B.  
TITLE Generation of ESTs from abiotic stressed potato tissue  
JOURNAL Unpublished (2003)

COMMENT Other ESTs: EST717289  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from TIGR via potato@tigr.org  
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

source

1. 967  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/tissue\_type="abiotic stress treated leaf and root tissue"  
/lab\_host="DH10B-Tona"  
/clone\_lib="potato abiotic stress cDNA library"  
/note="Vector: PCWVSport6.1; Site 1: EcORI; Site 2: NotI;  
supplier: Solanum tuberosum var. Kennebec plants were  
grown from cuttings on a 16hr light/8 hr dark cycle at 25  
C for 3-4 weeks. Abiotic stress conditions were applied to  
four separate sets of plants. Set 1 involved saturation at  
the soil with 150 mM NaCl and tissues were harvested at  
following application of the salt stress (leaves: 2hr,  
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).  
Set 2 were grown under the standard conditions and then  
were water stressed by withdrawal of further watering  
applications. Drought stressed plants were harvested after  
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d  
and 5d). Set 3 were grown under the standard conditions  
and then were cold stressed by placement at 4 C. Cold  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
and 4d. Set 4 were grown under the standard conditions and  
then were heat stressed by placement at 35 C. Heat  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d and 4d and heat-stressed roots were harvested at 6 hr,  
12 hr, 1 d, and 4d. RNA was isolated from all tissues and  
equal RNA from each tissue and stress was pooled to  
construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:

Pred. No.:	4.03e-138	Length:	967
Score:	1372.00	Matches:	264
Percent Similarity:	94.10%	Conservative:	23
Best Local Similarity:	86.56%	Mismatches:	18
Query Match:	84.53%	Indels:	0



DB: 14 Gaps: 0  
US-09-931-457A-31 (1-325) x CK271210 (1-967)

QY 1 MetAlaValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrPro 20  
DB 51 ATGGCCGGGGAAGAAAGATTGGAATGCCAAGAGATGTAAGTGAATGATTGTTACACTCTT 110  
QY 21 LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40  
DB 111 TTGGTATACCTGATATATGTTGATGGGTGTGTGACAGTGTGTGCTGCCAAGCTGCA 170  
QY 41 LeuMetGluProCysSerSerValIleAspArgIleGlyTyrSerMetIleAlaAspAla 60  
DB 171 AGCATGAGACCATGCTCTAGTGTGATGATAGATCGGTATAGTATGATTACAGATGCT 230  
QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80  
DB 231 GAGGAGAAAGGCTCTCATCAAAACCCGCGAGAGAGTGTCTCATGCAACCTAGAGTGAAC 290  
QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMet 100  
DB 291 ACTGGTGTAGGATTGGCAATTCATGCTGCTGCTAAAGGCTTACAACTCATCATACGATG 350  
QY 101 ProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120  
DB 351 CCTTCTTCATGATGCTTGAAGAGAAATATTCTGCGTCTTCTGCTGCTGAGTGTG 410  
QY 121 LeuThrAspProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluGluIleLeuAla 140  
DB 411 CTTACCGATCCAGCAAAAGGAGTGAAGGGTCTATTTCAAAGGCTGAAAGATTAAGGCC 470  
QY 141 LysThrProAsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHis 160  
DB 471 AAAACACCAACTCCTATATCTTCAAGCAATTTGAAAACCTGCTTAAACCAAGATACAC 530  
QY 161 TyrGluThrThrGlyProGluIleThrLysGlySerAspGlyLysIleAspAlaPheVal 180  
DB 531 TATGAGCACTGCTGCTGAGATCTGGAAGGCTCAAAATGGGAAGTGAATGCTTACTG 590  
QY 181 SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsn 200  
DB 591 TCTGGAATGGAACAGAGGACAGATACGCTTCAAGCAAGTATTGAGAGAGAGAGAAC 650  
QY 201 ProAsnLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLys 220  
DB 651 CCCAAGCTTAAGCTGTATGGCGTGAACCAATTGAAGTGTATCTTCTGCTGAAG 710  
QY 221 ProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240  
DB 711 CCTGTCACATAGATTCAGAGGAGTGTGCTGCTGCTTCAATCTGCTGCTTGAAGT 770  
QY 241 AsnLeuLeuAspGluValValGlnLysSerSerAspGluAlaIleGluThrAlaLysLeu 260  
DB 771 AACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 830  
QY 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla 280  
DB 831 CTGGCATGGAAGAGATGCTAGTGAATTCATCTGCTGCTGCTGCGCGCGCGCA 890  
QY 281 PheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSer 300  
DB 891 ATTAAGTGGTAAAGCGCCCTGAAATGCTGGAAGCTCATGTTGTTATTTTCCACG 950  
QY 301 PheGlyGluArgTyr 305  
DB 951 TTCGAGAGCGCATAT 965  
RESULT 5  
CB974516 910 bp mRNA linear EST 01-MAY-2003  
LOCUS CB974516  
DEFINITION CAB30005.1iib\_Fb\_P10 Cabernet Sauvignon Berry Stage I - CAB3 Vitis vinifera cDNA clone CAB30005.1iib\_Fb\_P10 5', mRNA sequence.

ACCESSION CB974516  
VERSION CB974516.1 GI:30297722  
KEYWORDS EST.  
SOURCE Vitis vinifera  
ORGANISM Vitis vinifera  
REFERENCE Bukaryota; Vitridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; Vitaceae; Vitis.  
1 (bases 1 to 910)  
Goes da Silva, F., Iandolino, A., Lim, H., Baek, J., Jones, K. and Cook, D.  
TITLE Expressed sequence tags from Vitis vinifera 'Cabernet Sauvignon' berries at various developmental stages  
JOURNAL Unpublished (2003)  
COMMENT Contact: Douglas Cook, PhD  
CAES Genome Facility  
UC Davis, Plant Pathology  
One Shields Ave, Davis, CA 95616, USA  
Tel: 530 754 6561  
Fax: 530 754 6617  
Email: drcoc@ucdavis.edu  
Seq primer: ACCGTACCGACATATGCC.  
FEATURES  
source  
1..910  
/organism="Vitis vinifera"  
/mol\_type="mRNA"  
/cultivar="Cabernet Sauvignon"  
/db\_xref="taxon:29760"  
/clone="CAB30005.1iib\_Fb\_P10"  
/sex="Hermaphrodite"  
/dev\_stage="Berry stage I"  
/lab\_host="DH5alpha"  
/clone\_lib="Cabernet Sauvignon Berry Stage I - CAB3"  
/note="Organ: Berry; Vector: pDNR; Site\_1: S11; Site\_2: S11; CAB3 is a cDNA library of Vitis vinifera 'Cabernet Sauvignon' Clone 8 berries. Samples were collected after berry set from field-grown vines during stage I of berry growth, 17 days after full bloom. The average berry size was 6 millimeters. Sampled vines were located at the University of California, Davis, Experimental Vineyard. cDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows: 5'-AAGCAGTGTATCAACGACAGATGCGCATTAAGCGCGG-3' and 5'-ATTCTAGAGCCGAGCGCGCGCATATG-dt(30)NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

ORIGIN  
Alignment Scores:  
Pred. No.: 1.27e-137 Length: 910  
Score: 1367.00 Matches: 268  
Percent Similarity: 95.68% Conservative: 20  
Best Local Similarity: 89.04% Mismatches: 12  
Query Match: 84.23% Indels: 1  
DB: 14 Gaps: 0  
US-09-931-457A-31 (1-325) x CB974516 (1-910)  
QY 8 ILeAlaLysAspValThrGluLeuIleGlyLysThrProLeuValTyrLeuAsnLysLeu 27  
DB 8 ATGGCCAAAGATGTTACAGATTAATTTGGAAACACCAATTAGTATATCTGAACATGTG 67  
QY 28 AlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluProCysSerSer 47  
DB 68 GTGATGGCTGTGATGCTGTATGCTGCAAGCTTGAATGATGAGAGCTGCTTACTG 127  
QY 48 ValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGluLysGlyLeuIleThr 67  
DB 128 GTCAAGACAGAGATCGGTATGATGATGATGATGATGATGATGATGATGATGATGAT 187  
QY 68 ProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIleGlyLeuAlaPhe 87  
DB 188 CCAGGAGAGAGTGTCTCTCATTTGAGCTTACCAAGTGTGTAATCTGCGATAGGGTTCATT 247



QY 88 MetAlaAlaAlaArgGlyTyrIleuLeuIleIleThrMetProAlaSerMetSerLeuGlu 107  
 Db 248 ATGGCCGGCAGCTAAGGCTTATTAAGCTCATCATTAACAATGCTCTTCGATGAGCTTGAAG 307  
 QY 108 ArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspProAlaAlaGly 127  
 Db 308 AGAAGAATTAATTCCTCCAGCTTTGGGGCTGAATTGGTTCTTCACAGATCCAGCCAGAGGC 367  
 QY 128 MetIysGlyAlaAlaValGlnIleAlaGluGluIleLeuAlaIleThrProAlaAlaTyrIle 147  
 Db 368 ATGAAGGGGGCTGTTCAAGAGCTGAAGAGATAGTGGCAAGACACCTAATTCTATATT 427  
 QY 148 LeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThrGlyProGlu 167  
 Db 428 CTTCAGCAATTGAAACCTCCCAATCCGAAGATTATGAAACCACTGACCCGAG 487  
 QY 168 IleTyrIleGlySerAspGlyLysIleAspAlaPheValSerGlyIleGlyThrGlyGly 187  
 Db 488 ATCTGAGAGAGGCACTGTGGGAAGTGTATGCTTTGCTCTGGGATAGGAGCTGAGGT 547  
 QY 188 ThrIleThrGlyAlaGlyLysTyrLeuLysGlyGlnAsnProAlaIleLysLeuIleGly 207  
 Db 548 ACCATAACAGGTGAGGAGGAGTCTCAAGAGCAAAATTTGACCTTAAGCTGTATGCT 607  
 QY 208 ValGluProValGluSerProValLeuSerGlyGlyLysProGlyProHisIleGln 227  
 Db 608 GTAGAACCTGTGGAAGTGCAGTACTGTCTGAGGAAAACTGGCCGCAATAGATCCAA 667  
 QY 228 GlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeuLeuAspGluVal 247  
 Db 668 GGAATGGCGCTGCTTCAATCCCTGAGATTAGATGCAATTTGCTGATGATGTT 727  
 QY 248 GlnIleSerSerAspGluAlaIleGluThrAlaIleLeuAlaLeuGlyGluLeu 267  
 Db 728 CAAGTTCAAGTGAAGAGCTGTGAAGCTGCCAAGCTTCTTGCAATGAAGAGTTTG 787  
 QY 268 PheValGlyIleSerSerGlyAlaAlaAlaAlaAlaPheGlnIleAlaLysArgPro 287  
 Db 788 CTGCTGGGATATCATCTGTGTGTCTGACGCTGCTGCAATAAAGTTAGCAAGAGGCCA 847  
 QY 288 GluAsnAlaGlyLysLeuIle-ValAlaValPheProSerPheGlyGluArgTyrLeu 307  
 Db 848 GAAATGCTGGAAATCTCATTTGGGATGCTTTCCCAAGCTTTGAGAGCGTTATCTGTC 907  
 QY 307 r 307  
 Db 908 c 908

RESULT 6  
 LOCUS CK266154 937 bp mRNA linear EST 12-DEC-2003  
 DEFINITION EST712232 potato abiotic stress cDNA library Solanum tuberosum cDNA  
 clone POAC308 5' end, mRNA sequence.  
 ACCESSION CK266154  
 VERSION CK266154.1 GI:39823132  
 KEYWORDS EST.  
 SOURCE Solanum tuberosum (potato)  
 ORGANISM Solanum tuberosum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asterids; Lamiales; Solanales; Solanaceae; Solanum.  
 1 (bases 1 to 937)  
 Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.  
 Generation of ESTs from abiotic stressed potato tissue  
 Unpublished (2003)  
 Other ESTs: EST712233  
 Contact: Robin Buell  
 The Institute for Genomic Research  
 9712 Medical Center Dr, Rockville, MD 20850, USA  
 Email: potato-array@tigr.org  
 Clones can be requested from TIGR via potato@tigr.org  
 Seq primer: ATT TAG GTG ACA CTA TAG.

# FEATURES

## source

Location/Qualifiers  
 1. 937  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cultivar="Kennebec"  
 /db\_xref="taxon:4113"  
 /clone="POAC308"  
 /tissue\_type="abiotic stress treated leaf and root tissue"  
 /lab\_host="DH10B-Tona"  
 /clone\_lib="potato abiotic stress cDNA library"  
 /notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
 supplier: Solanum tuberosum var. Kennebec plants were  
 grown from cuttings on a 16hr light/8 hr dark cycle at 25  
 C for 3-4 weeks. Abiotic stress conditions were applied to  
 four separate sets of plants. Set 1 involved saturation of  
 the soil with 150 mM NaCl and tissues were harvested at  
 following application of the salt stress (leaves: 2hr,  
 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).  
 Set 2 were grown under the standard conditions and then  
 were water stressed by withdrawal of further watering  
 applications. Drought stressed plants were harvested after  
 cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d  
 and 5d). Set 3 were grown under the standard conditions  
 and then were cold stressed by placement at 4 C. Cold  
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
 and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
 2d. Set 4 were grown under the standard conditions and  
 then were heat stressed by placement at 35 C. Heat  
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
 2d and 4d and heat-stressed roots were harvested at 6 hr,  
 12 hr, 1 d, and 4d. RNA was isolated from all tissues and  
 equal RNA from each tissue and stress was pooled to  
 construct the cDNA library. RNA sample."

## ORIGIN

### Alignment Scores:

Pred. No.: 2.83e-137 Length: 937  
 Score: 1364.00 Matches: 263  
 Percent Similarity: 93.79% Conservative: 24  
 Best Local Similarity: 85.95% Mismatch: 19  
 Query Match: 84.04% Indels: 0  
 DB: 14 Gaps: 0

US-09-931-457A-31 (1-325) x CK266154 (1-937)

QY 1 MetAlaValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrPro 20  
 Db 20 ATGGCCGGGGAAGATTGGAATGCCAAGATGTAAGTGAATGATGTAACACTCCT 79  
 QY 21 LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40  
 Db 80 TTGCTATACCTGAATATATGTGTGATGGGTGTGTGACAGTGTCTGCCAAGCTGAA 139  
 QY 41 LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60  
 Db 140 AGCATGAGGCAATGCTCTAGTGTGAAGATAGGATCGGTTATGATGATTAAGATGCT 199  
 QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80  
 Db 200 GAGGAGAGGGTCTCATCAACCCGGGAGAGTGTCTCATCAACCTAAGAGTGAAC 259  
 QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaArgGlyTyrIleLeuIleIleThrMet 100  
 Db 260 ACTGGTGAAGATTGGCATTCAAGCTGCTGCTTAAAGGCTCAAACTCATCATGAGATG 319  
 QY 101 ProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120  
 Db 320 CCTTCTCAATGAGCTTGAAGAGAAATTAATCTGCGTCTTCGGTCTGAGTTGTTG 379  
 QY 121 LeuThrAspProAlaLysGlyMetIysGlyAlaValGlnLysAlaGluGluIleLeuAla 140  
 Db 380 CTACCAATCCAGCAAAAGGAGTGAAGGCTGCTATTTCAGAGGCTGAAGAGATTAAGGCC 439

QY 141 LysThrProAsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHis 160  
 DB 440 AAAACACCCCACTCTATATCTCTACGCAATTGAAAACCTGCTACCAACCAAGATACAC 499  
 QY 161 TyrGluThrThrGlyProGluIleTfplyGlySerAspGlyLysIleAspAlaPheVal 180  
 DB 500 TATGAGACCACTGCTCTGAGATCTGGAAGGCTCAATGGGAAGTGGATGCTCTAGTC 559  
 QY 181 SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsn 200  
 DB 560 TCTGAATTGGAACAGAGGACGATGCTGTTCAAGCAAGTATTGAGAGAGCAAGAAC 619  
 QY 201 ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyGlyLys 220  
 DB 620 CCCAAGCTTAAGCTTATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 679  
 QY 221 ProGlyProHisIleGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 240  
 DB 680 CCTGCTCCACATAGATTCAGGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 739  
 QY 241 AsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260  
 DB 740 AACCTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 799  
 QY 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla 280  
 DB 800 CTGGCATTTGAGGAGGAGGATGCTAGTGGGATTTTCATCTGCTGCTGCTGCTGCTGCT 859  
 QY 281 PheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSer 300  
 DB 860 ATTAAGTCGCTAAGCGCCCTGAAATGCTGGGAAAGCTCATTTGTGATTTTCCACAGC 919

QY 301 PheGlyGluArgTyrLeu 306  
 DB 920 TTCGAGAGCGCATATCTT 937

RESULT 7  
 CK259489  
 LOCUS EST705567 potato abiotic stress cDNA library Solanum tuberosum cDNA  
 DEFINITION  
 CK259489  
 VERSION CK259489.1 GI:39816467  
 KEYWORDS  
 SOURCE Solanum tuberosum (potato)  
 ORGANISM Solanum tuberosum  
 BUKARYOTA; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; Lamiales; Solanales; Solanaceae; Solanum.  
 1 (bases 1 to 881)  
 Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.  
 Generation of ESTs from abiotic stressed potato tissue  
 Unpublished (2003)  
 Other ESTs: EST705565 EST705566  
 Contact: Robin Buell  
 The Institute for Genomic Research  
 9712 Medical Center Dr, Rockville, MD 20850, USA  
 Email: potatocenter@tigr.org  
 Clones can be requested from TIGR via potatocenter@tigr.org  
 Seq primer: CAG GAA ACA GCT ATG ACC.  
 Location/Qualifiers  
 1. 881  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /culturivar="Kennebec"  
 /db\_xref="taxon:4113"  
 /clone="POAB060"  
 /issue\_type="abiotic stress treated leaf and root tissue"  
 /lab\_host="DH10B-Tona"  
 /clone\_lib="potato abiotic stress cDNA library"  
 /notes="vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
 supplier: Solanum tuberosum var. Kennebec plants were  
 grown from cuttings on a 16hr light/8 hr dark cycle at 25

C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, 1d, 2d, and 4d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4 d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4 d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 8.89e-132 Length: 881  
 Score: 1313.00 Matches: 253  
 Percent Similarity: 94.52% Conservative: 23  
 Best Local Similarity: 86.64% Mismatches: 16  
 Query Match: 80.90% Indels: 0  
 DB: 14 Gaps: 0

US-09-931-457A-31 (1-325) x CK259489 (1-881)

QY 4 GluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrProLeuValTyr 23  
 DB 2 GAAAGATTGGAATTGCCAAGATGTAAGTGAATTGATTGTAACACTCTTGGTATAC 61  
 QY 24 LeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGlu 43  
 DB 62 CTGATAATGTTGTGATGGGTGTGTGCACGTTGCTGCCAAGCTGGAAGCATGGAG 121  
 QY 44 ProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluLys 63  
 DB 122 CCATGCTCTAGTGTTAAGATAGATCGTTATAGTATGATTAACAGATGCTGAGAGAG 181  
 QY 64 GlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIle 83  
 DB 182 GGTCTCATCAAAACCCGCGAGAGTGTCTCATCGAACCTACGAGTGAACACTGGTGA 241  
 QY 84 GlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProAlaSer 103  
 DB 242 GGATTGGCATTTCATGCTGCTGCTAAAGGCTACAACTCATCATGATGCTCTTCA 301  
 QY 104 MetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAsp 123  
 DB 302 ATGAGCTTGAAGAGAAATTAATTTCTGCGTCTTCCGTGCTGAGTTGGTCTTACCAT 361  
 QY 124 ProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluLysIleLeuAlaLysThrPro 143  
 DB 362 CCAGCAAAAGGATGAGGGTGTCTAATTCAAAGGCTGAGAGATTAAGGCCAACAACACC 421  
 QY 144 AsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThr 163  
 DB 422 AACTCTATATTTCTTCAAGCAATTTGAAAACCTGCTAACCCAAAGATACATAGAGACC 481  
 QY 164 ThrGlyProGluIleTfplyGlySerAspGlyLysIleAspAlaPheValSerGlyIle 183  
 DB 482 ACTGCTCTGAGATCTGGAAGGCTCAAAATGGGAAGTGAATGCTCTAGTCTCTGAAATT 541  
 QY 184 GlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsnProAsnIle 203  
 DB 542 GGAACAGAGGACGATAGTCTGCTCAGGCAAGTATTGAGAGAGCAAGACCCCAAGCTT 601  
 QY 204 LysLeuIleGlyValGluProValGluSerProValLeuSerGlyLysProGlyPro 223

```

|||||
Db 602 AAGCTGATATGCGGTGAACCAAGTGAAGTGTATCTTCTGTTGGAAGCCTGTCCA 661
Qy 224 HisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValIleGluValAsnIleu 243
Db 662 CATAGATTGAGGGGATGGTGTGTTTCTTCTGTTTGAAGTTAACTTATT 721
Qy 244 AspGluValAlaGlnIleSerSerAspGluAlaIleGluThrAlaLysIleuAlaIleu 263
Db 722 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 781
Qy 264 LysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAlaAlaPheGlnIle 283
Db 782 AAGGAGGATGCTAGTGGGATTTCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 841
Qy 284 AlaLysArgProGluAsnAlaGlyLysIleuIleVal 295
Db 842 GCTAAGCGCCCTGAAATATGCTGGGAGCTCATTTGT 877

```

```

RESULT 8
LOCUS CK260012 907 bp mRNA linear EST 12-DEC-2003
DEFINITION EST706090 potato abiotic stress cDNA library Solanum tuberosum cDNA
clone POAB228 5' end, mRNA sequence.
ACCESSION CK260012
VERSION CK260012.1 GI:39816990
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 907)
AUTHORS Buehl, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
TITLE Generation of ESTs from abiotic stressed potato tissue
JOURNAL Unpublished (2003)
COMMENT Other ESTs: EST706088 EST706089 EST706091
Contact: Robin Buehl
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: CAG GAA ACA GCT ATG ACC.
Location/Qualifiers
1..907
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POAB228"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-Tona"
/clone_lib="potato abiotic stress cDNA library"
/notes="Vector: pCMVSPORT6.1, Site 1: EcoRI, Site 2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
following application of the salt stress (leaves: 2hr,
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, 1d,
2d). Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d and 4d and heat-stressed roots were harvested at 6 hr,

```

```

FEATURES
source

```

12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

```

ORIGIN
Alignment Scores:
Pred. No.: 1.54e-131 Length: 907
Score: 1311.00 Matches: 253
Percent Similarity: 94.18% Conservative: 22
Best Local Similarity: 86.64% Mismatches: 17
Query Match: 80.78% Indels: 0
DB: 14 Gaps: 0

```

US-09-931-457A-31 (1-325) x CK260012 (1-907)

```

Qy 1 MetAlaValAlaGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrPro 20
Db 31 ATGGCGGGGAAAGATTGGAATTGCCAAGATGTAACGAATTGATTGTAACACTCCT 90
Qy 21 LeuValIleThrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40
Db 91 TTGGTATACCTGAATATATGTTGTGATGGGTGTGTGACAGTGTGTGCTGCCAAGCTGAA 150
Qy 41 LeuMetGluProCysSerSerValLysAspArgIleGlyIleThrSerMetIleAlaAspAla 60
Db 151 AGCATGAGCCATGCTCTAGTGTAAAGATAGATCGGTATAGTATGATTAAGATGCT 210
Qy 61 GluGluIleGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80
Db 211 GAGGAGAAAGGCTCATCAACCCGGCAGAGTGTCTCATCGAAGCTCAAGTGAAGAAC 270
Qy 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyIleThrLysLeuIleIleThrMet 100
Db 271 ACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 330
Qy 101 ProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120
Db 331 CCTTCTTCAATGATCTTGAAGAGAAATATCTGCGTGTCTTGGTGTGATGATGATGATG 390
Qy 121 LeuThrAspProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluIleLeuAla 140
Db 391 CTTACCGATCCACCAAGGATGAGGCTGCTATTTCAGAGGCTCAAGATTAAGGCC 450
Qy 141 LysThrProAsnAlaTyrIleLeuGlnIlePheGluAsnProAlaAsnProLysValHis 160
Db 451 AAACACCCCACTCTATATCTTCAAGCAATTGAAACCCCTGCTAACCCCAAGATACAC 510
Qy 161 TyrGluThrThrGlyProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheVal 180
Db 511 TATGAGACCACTGCTCTGAGATCTGGAAGGCTCAATGCGGAAGTGAATGCTCTAGTC 570
Qy 181 SerGlyIleGlyThrGlyIleThrIleThrGlyAlaGlyLysIleLeuLysGluGlnAsn 200
Db 571 TCTGGAATTGGAACAGAGGACAGATTAAGCTGTTCAAGCAAGTATTGAGAGAGCAGAAC 630
Qy 201 ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLys 220
Db 631 CCCAAGCTTAAGCTGATGCGCTGGAACCAAGTTGAAAGTGTATCTTCTGTTGAAAG 690
Qy 221 ProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240
Db 691 CCGTGTCCATTAAGATTCAAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 750
Qy 241 AsnLeuLeuAspGluValAlaGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260
Db 751 AACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 810
Qy 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla 280
Db 811 CTGCAATTGAAGAGGATGCTAGTGGGAAATTCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 870
Qy 281 PheGlnIleAlaLysArgProGluAsnAlaGlyLys 292

```



Db 871 ATTAAGTCCTAAGCGCCCTGAATGCTGGAG 906

## RESULT 9

CK259857

## LOCUS

CK259857

## DEFINITION

EST705935 potato abiotic stress cDNA library Solanum tuberosum cDNA clone POAB178 5' end, mRNA sequence.

## ACCESSION

CK259857

## VERSION

CK259857.1 GI:39816835

## KEYWORDS

EST.

## SOURCE

Solanum tuberosum (potato)

## ORGANISM

Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.

## REFERENCE

1 (bases 1 to 920)

Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.

## AUTHORS

Generation of ESTs from abiotic stressed potato tissue

## TITLE

Unpublished (2003)

## JOURNAL

Other ESTs: EST705934 EST705936 EST705937

## COMMENT

Contact: Robin Buell

## REFERENCE

The Institute for Genomic Research

## AUTHORS

9712 Medical Center Dr, Rockville, MD 20850, USA

## TITLE

Email: potato-array@tigr.org

## JOURNAL

Clones can be requested from TIGR via potato@tigr.org

## COMMENT

Seq primer: ATT TAG GTG ACA CTA TAG.

## FEATURES

Location/Qualifiers

## SOURCE

1..920

## ORGANISM

/organism="Solanum tuberosum"

## MOLTYPE

/mol\_type="mRNA"

## CULTIVAR

/cultivar="Kennebec"

## DBXREF

/db\_xref="taxon:4113"

## CLONE

/clone="POAB178"

## TISSUE

/tissue\_type="abiotic stress treated leaf and root tissue"

## LAB\_HOST

/lab\_host="DH10B-Tona"

## CLONE\_LIB

/clone\_lib="potato abiotic stress cDNA library"

## NOTE

/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: Solanum tuberosum var. Kennebec plants were grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

## ORIGIN

## Alignment Scores:

Pred. No.:

1.58e-131

Length:

920

Score:

1311.00

Matches:

253

Conservative:

22

Version:

17

Mismatch:

0

Indels:

0

Gaps:

0

DB:

14

US-09-931-457a-31 (1-325) x CK259857 (1-920)

1 MetAlaValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyIleThrPro

20

21 MetAlaValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyIleThrPro

20

22 MetAlaValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyIleThrPro

20

23 MetAlaValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyIleThrPro

20

24 MetAlaValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyIleThrPro

20

25 MetAlaValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyIleThrPro

20

26 MetAlaValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyIleThrPro

20

27 MetAlaValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyIleThrPro

20

28 MetAlaValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyIleThrPro

20

29 MetAlaValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyIleThrPro

20

30 MetAlaValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyIleThrPro

20



TITLE	Generation of ESTs from abiotic stressed potato tissue
JOURNAL	Unpublished (2003)
COMMENT	Other_ESTs: EST705565 EST705567

## FEATURES

### Location/Qualifiers

/tissue\_type="Abiotic stress treated leaf and root tissue  
/lab\_host="DH10B-Tona"  
/clone\_lib="potato abiotic stress cDNA library"  
/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: Solanum tuberosum var. Kennebec plants were  
grown from cuttings on a 16hr light/8 hr dark cycle at 25  
C for 3-4 weeks. Abiotic stress conditions were applied to  
four separate sets of plants. Set 1 involved saturation of  
the soil with 150 mM NaCl and tissues were harvested at  
following application of the salt stress (leaves: 2hr,  
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).  
Set 2 were grown under the standard conditions and then  
were water stressed by withdrawal of further watering  
applications. Drought stressed plants were harvested after  
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d  
and 5d). Set 3 were grown under the standard conditions  
and then were cold stressed by placement at 4 C. Cold  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
and 4d. Set 4 were grown under the standard conditions and  
then were heat stressed by placement at 35 C. Heat  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d and 4d and heat-stressed roots were harvested at 6 hr,  
12 hr, 1 d, and 4d. RNA was isolated from all tissues and  
equal RNA from each tissue and stress was pooled to  
construct the cDNA library. RNA sample."

QY	103	SerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThr	122
DB	301	TCAATGAGCTTGAGAGAGAATATTCTGCGTCTTCCGTGCTGAGTGGTGCTTACC	360
QY	123	AspProAlaIysGlyMetLysGlyAlaValGlnLysAlaGluGluIleLeuAlaIysThr	142
DB	361	GATCCAGCAAAAGCGATGAAGGCTCTATTCAAAGCTGAAGAGATAAAGCCAAACA	420
QY	143	ProAsnAlaIyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGlu	162
DB	421	CCCAACTCCTATTTCTTCTCAGCAATTTGAAACCCCTGCTAACCCAAAGATACACTATGAG	480
QY	163	ThrThrGlyProGluIleTrpLysGlySerAspGlyLysIleAspAlaPheValSerGly	182
DB	481	ACCACCTGGCTCCTAGATCTGGAAGGCTCAATGGGAAAGTGATGCTCTAGTCTTGGA	540
QY	183	IleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsnProAsn	202
DB	541	ATTGAACAGAGAGCAAGATAACTGTTCTAGGCAAGTATTTGAGAGACAGAACCCCAAC	600
QY	203	IleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyGlySerProGly	222
DB	601	GTTAAGCTGATAGCGGTGGAACCAAGTTGAAGTCTATCCTTCTGCTGGAAGCCTGGT	660
QY	223	ProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeu	242
DB	661	CCACATAAGATTCAAGGGAATTGGTCTGCTTCAATTCCTGGTGTGGGAAGTTAACCTT	720
QY	243	LeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeuLeuAla	262
DB	721	ATTGATGATGTAGTTCAGGTTTCAAGTGAAGAAATCCATAGAAATGGCTAAGCTTCTGCA	780
QY	263	LeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAlaAlaPheGln	282
DB	781	TTGAAGGAAGGATTGCTAGTGGGAATTCATCTGCTGCTGCGGCGGCAATTAA	840
QY	283	IleAlaLysArgProGluAsnAlaGlyLysLeu	293
DB	841	GTCGCTAAGCGCCCTGAAAAATGCTGGGAAGCTC	873

RESULT 11

LOCUS	917 bp	mRNA	linear	EST 12-DEC-2003
DEFINITION	CK259858			
	EST705936	potato abiotic stress cDNA library	Solanum tuberosum	cDNA
	clone POAB178 5' end,	mRNA sequence.		

ACCESSION	CK259858	GI:39816836
VERSION	CK259858.1	
KEYWORDS	EST.	
SOURCE	Solanum tuberosum (potato)	
ORGANISM	Solanum tuberosum	

REFERENCE	AUTHORS	JOURNAL	COMMENT
1 (bases 1 to 917)	Buelli, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B	Generation of ESTs from abiotic stressed potato tissue	
	Unpublished (2003)		
	Other_ESTs: EST705934 EST705935 EST705937		

FEATURES	Location/Qualifiers
source	1. .917

/lab host="DH10B-Tona"  
/clone lib="potato abiotic stress cDNA library"  
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: Solanum tuberosum var. Kennebec plants were  
grown from cuttings on a 16hr light/8 hr dark cycle at 25  
C for 3-4 weeks. Abiotic stress conditions were applied to  
four separate sets of plants. Set 1 involved saturation of  
the soil with 150 mM NaCl and tissues were harvested at  
following application of the salt stress (leaves: 2hr,  
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).  
Set 2 were grown under the standard conditions and then  
were water stressed by withdrawal of further watering  
applications. Drought stressed plants were harvested after  
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d  
and 5d). Set 3 were grown under the standard conditions  
and then were cold stressed by placement at 4 C. Cold  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d. Set 4 were grown under the standard conditions and  
then were heat stressed by placement at 35 C. Heat  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d and 4d and heat-stressed roots were harvested at 6 hr,  
12 hr, 1 d, and 4d. RNA was isolated from all tissues and  
equal RNA from each tissue and stress was pooled to  
construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:

Pred. No.: 5.47e-131 Length: 917  
Score: 1306.00 Matches: 252  
Percent Similarity: 94.16% Conservative: 22  
Best Local Similarity: 86.60% Mismatches: 17  
Query Match: 80.47% Indels: 0  
DB: 14 Gaps: 0

US-09-931-457A-31 (1-325) x CK259858 (1-917)

QY 1 MetAlaValGluArgSerGlyIleAlaIysAspValThrGluLeuIleGlyLysThrPro 20  
DB 45 ATGGCGGGGAAAGATTGGAAATGCCAAGATGTAAGTAATGATTGTAACACTCTCT 104  
QY 21 LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40  
DB 105 TTGGTATACCTGAATATATGTTGTGATGGGTGGTGTGACAGTGTGTGCCAAGCTGGAA 164  
QY 41 LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60  
DB 165 AGCATGAGCCATGCTCTAGTGTTAAGATAGATCGCTTATAGATGATTAACAGATGCT 224  
QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80  
DB 225 GAGGAGAAAGGCTCATCAAAACCCGGCAGAGTGTCTCATCGAACCTACAGAGTGAAC 284  
QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMet 100  
DB 285 ACTGGTGTAGATGGATTCATTCATGCTGCTGCTTAAAGGCTACAACTCATCATTCAGATG 344  
QY 101 ProAlaSerMetSerLeuGluArgGluArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120  
DB 345 CCTCTTCAATGAGTCTTGAGAGAGAATATTTCTGCTGCTTTCGCTGCTGAGTGTG 404  
QY 121 LeuThrAspProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluGluIleLeuAla 140  
DB 405 CTTACCGATCCAGCAAAAGGATGAAGGGTCTATTTCAAAGGCTGAAGAGATAAAGGCC 464  
QY 141 LysThrProAsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHis 160  
DB 465 AAAACACCCCACTCTATATCTTCAGCAATTGAAACCCCTGCTAACCCCAAGATACAC 524  
QY 161 TyrGluThrThrGlyProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheVal 180  
DB 525 TATGAGACCACTGCTCTGAGATCTGGAAGGCTCAATGGAAGGATGCTCTAGTCT 584

QY 181 SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsn 200  
DB 585 TCTGGAATGGAACAGAGAGCAGATTAAGTGTTCAGGCAAGTATTGAGAGACAGAAC 644  
QY 201 ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyGlyLys 220  
DB 645 CCCAAGCTTAAGCTGTATGCGGTGAACCAAGTTGAAGTGTCTCTTCTGTTGAAG 704  
QY 221 ProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240  
DB 705 CTTGTCACATTAAGATTCAAGGAGATTGCTGCTGTTTCAATTCCTGTTTGAAGTT 764  
QY 241 AsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260  
DB 765 AACCTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 824  
QY 261 LeuAlaLeuLysGluGluGluPheValGlyIleSerSerGlyAlaAlaAlaAlaAla 280  
DB 825 CTGGATTGAAGGAAGATTGCTAGTGGGAATTCATCTGCTGCTGCTGCGCGCGCA 884  
QY 281 PheGlnIleAlaLysArgProGluAsnAlaGly 291  
DB 885 ATTAAGTCGCTAAGCGCCCTGAATAATGCTGGG 917

RESULT 12

CK279506

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 935)

Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.

Generation of ESTs from abiotic stressed potato tissue

Unpublished (2003)

Other ESTs: EST725585

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from TIGR via potato@tigr.org

Seq primer: ATT TAG GTG ACA CTA TAG.

Location/Qualifiers

1. 935

/organism="Solanum tuberosum"

/mol\_type="mRNA"

/cultivar="Kennebec"

/db\_xref="taxon:4113"

/clone="POAEC59"

/tissue\_type="abiotic stress treated leaf and root tissue"

/lab\_host="DH10B-Tona"

/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: Solanum tuberosum var. Kennebec plants were  
grown from cuttings on a 16hr light/8 hr dark cycle at 25  
C for 3-4 weeks. Abiotic stress conditions were applied to  
four separate sets of plants. Set 1 involved saturation of  
the soil with 150 mM NaCl and tissues were harvested at  
following application of the salt stress (leaves: 2hr,  
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).  
Set 2 were grown under the standard conditions and then  
were water stressed by withdrawal of further watering  
applications. Drought stressed plants were harvested after  
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d  
and 5d). Set 3 were grown under the standard conditions  
and then were cold stressed by placement at 4 C. Cold  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,

and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

## ORIGIN

## Alignment Scores:

Pred. No.:	2,41e-129	Length:	935
Score:	1291.00	Matches:	254
Percent Similarity:	93.27%	Conservative:	23
Best Local Similarity:	85.52%	Mismatches:	20
Query Match:	79.54%	Indels:	1
DB:	14	Gaps:	0

US-09-931-457A-31 (1-325) x CK279506 (1-935)

```

QY 1 MetAlaValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrPro 20
DB 46 ATGGCGGGGAAAGATTGGAATTGCCAAGATGTAAGTGAATTGATTGTAACACTCCT 105
QY 21 LeuValThrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40
DB 106 TTGGTATACCTGAATATATGTTGTGATGGGTGGTGTGCAAGTGTGCTGCAAGCTGGA 165
QY 41 LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60
DB 166 AGCATGAGCCATGCTCTACTGTTAAGATAGATCGGTTATAGATGATTAACAGATGCT 225
QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80
DB 226 GAGGAGAAGGGTCTCATCAAAACCGCGCAGAGTGTCTCATGAACTCAAGTGAAGAAC 285
QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMet 100
DB 286 ACTGTGTAGATGTCATTCATGCTGCTGCTTAAGGCTTAACAACTCATTAAGATG 345
QY 101 ProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyValGluLeuVal 120
DB 346 CCTTCTCAATGATCTTGAGAGAAGAATTAATCTGCTGCTTTCGCTGCTGAGTGTG 405
QY 121 LeuThrAspProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluGluIleLeuAla 140
DB 406 CTTACCGATCCAGCAAAAGGATGAAGGTCGTATTTCAAAAGCTGAAGAGATAAAGGCC 465
QY 141 LysThrProAsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHis 160
DB 466 AAAACACCCCACTCCTATATCTTCAAGCAATTGAAAACCTGCTAAACCAAGATACAC 525
QY 161 TyrGluThrThrGlyProGluIleTryptLysGlySerAspGlyLysIleAspAlaPheVal 180
DB 526 TATGAGACCACTGCTCTGAGATCTGAAAAGGCTCAATGGGAAGTGAATGCTCTAGTC 585
QY 181 SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsn 200
DB 586 TCTGAATTGGAACAGAGGACAGATACTGTTCAAGCAAGTATTTGAGAGACAGAAC 645
QY 201 ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLys 220
DB 646 CCCAAGCTTAAGCTGATGCGCTGGAACCAAGTGAAGTGTCTTCTGCTGGAAG 705
QY 221 ProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240
DB 706 CTTGTTCACATAGATTCAAGGGGATTTGCTGCTTTCATTCCTGCTGTTTGAATT- 764
QY 241 AsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260
DB 765 AACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 824
QY 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla 280

```

```

DB 825 CTGGCATTTGAAGAAAGATTCTAGTGGGAATTTCTCTGTGCTGCTGCGCGCGCA 884
QY 281 PheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaVal 297
DB 885 ATTAAGTCTGCTAGCCCTGTAATAATCTGGAAGCTCATTTGTTATT 935

```

## RESULT 13

CG848066

## DEFINITION

ZMMBB0319111.1 r ZMMBB Zea mays subsp. mays genomic clone

## ACCESSION

CG848066

## VERSION

CG848066.1

## KEYWORDS

GSS.

## SOURCE

Zea mays subsp. mays (maize)

## ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

## REFERENCE

Yu, Y., Kim, H.R., Hatfield, J., Soderlund, C., Bharti, A.K., Messing, J. and Wing, R.

## AUTHORS

Sequencing of the maize genome

## TITLE

Unpublished (2003)

## JOURNAL

Contact: Rod Wing

## COMMENT

Arizona Genomics Institute

Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ

85721-0088, USA

Tel: 520 626 3967

Fax: 520 621 9288

Email: http://genome.arizona.edu

PCR Primers

FORWARD: T7

BACKWARD: M13r

Plate: 0319

Seq primer: M13r

Class: BAC ends.

Location/Qualifiers

1. 934

/organism="Zea mays subsp. mays"

/mol\_type="genomic DNA"

/cultivar="B73"

/sub\_species="mays"

/db\_xref="taxon:4578"

/clone="ZMMBB0319111"

/lab\_host="DH10B"

/note="Vector: pBelBAC11; Site\_1: HindIII; Site\_2: HindIII; Zea mays L. sp. mays"

## ORIGIN

## Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-931-457A-31 (1-325) x CG848066 (1-934)

```

QY 8 IleAlaLysAspValThrGluLeuIleGlyLysThrProLeuValTyrLeuAsnLysLeu 27
DB 72 ATGGCAAGATGTCATCAATGATCGCAAAAGCGCACTGATATATCGAACAAGTT 131
QY 28 AlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluProCysSerSer 47
DB 132 GTTGAAGGATGTGTGCTCCATGCTGCTAAGTTGAGGCAATGAAACCTGCTCAAGT 191
QY 48 ValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGluLysGlyLeuIleThr 67

```







QY 206 IleglyValGluProValGluSerProValLeuSerGlyGlyProGlyProHisLys 225  
 DB 601 ATTGGTATTGAGCCTTCTGAAAGTAACACTCTCCGGTGGAAAACCTGGTCCACATAAG 660  
 QY 226 IleginglyIleglyValGlyPheIleProGlyValLeuGluValAsnLeuLeuAspGlu 245  
 DB 661 ATCCAGGGAATCGGCGAGGATTTGTCACAGAACTGGATAGCATATTTCTGATGAA 720  
 QY 246 ValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeuLeuAlaLeuLysGlu 265  
 DB 721 GTAATTGAGATCAAGTATGATGAGCTGTGAGACAGCAAAACAGTGGCTGTCAAGAA 780  
 QY 266 GlyLeuPheValGlyIleSerSerGlyValAlaAlaAlaAlaPheGlnIleAlaLys 285  
 DB 781 GGATTACTGTCGGAATCTCTCTGAGCGACCGCTGCTGCTGCAATAAGGTTGCCAAA 840  
 QY 286 ArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSerPheGlyGluArgTyr 305  
 DB 841 AGACCAAGAGATGCTGGAAGCTGATAGTGTGTGTTCCGAGCTTCGGCGAGAGGTAC 900  
 QY 306 LeuSerSerValLeuPheGluSerValArgArgGluAlaGluSerMetThrPheGluPro 325  
 DB 901 CTTTCATCTGTCTCTATCACTGATGATGAGAGAAATGCGAGACATGCAAGCTGAGCCA 960

RESULT 15  
 CK265667 886 bp mRNA linear EST 12-DEC-2003  
 LOCUS EST711745 potato abiotic stress cDNA library Solanum tuberosum cDNA  
 DEFINITION clone POAC011 5' end, mRNA sequence.  
 CK265667  
 VERSION CK265667.1 GI:39822645  
 EST.  
 KEYWORDS Solanum tuberosum (potato)  
 SOURCE Solanum tuberosum  
 ORGANISM Solanum tuberosum; Streptophyta; Embryophyta; Tracheophyta;  
 Eukaryota; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicots;  
 Spermatophyta; Lamiales; Solanales; Solanaceae; Solanum.  
 1 (bases 1 to 886)  
 Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.  
 Generation of ESTs from abiotic stressed potato tissue  
 Unpublished (2003)  
 Other ESTs: EST711746  
 Contact: Robin Buell  
 The Institute for Genomic Research  
 9712 Medical Center Dr, Rockville, MD 20850, USA  
 Email: potato-array@tigr.org  
 Clones can be requested from TIGR via potato@tigr.org  
 Seq primer: ATT TAG GTG ACA CTA TAG.  
 Location/Qualifiers  
 1. 886  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cultivar="Kennebec"  
 /db\_xref="taxon:4113"  
 /clone="POAC011"  
 /tissue\_type="abiotic stress treated leaf and root tissue"  
 /lab\_host="DH10B-Tona"  
 /clone\_lib="potato abiotic stress cDNA library"  
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
 supplier: Solanum tuberosum var. Kennebec plants were  
 grown from cuttings on a 16hr light/8 hr dark cycle at 25  
 C for 3-4 weeks. Abiotic stress conditions were applied to  
 four separate sets of plants. Set 1 involved saturation of  
 the soil with 150 mM NaCl and tissues were harvested at  
 following application of the salt stress (leaves: 2hr,  
 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).  
 Set 2 were grown under the standard conditions and then  
 were water stressed by withdrawal of further watering  
 applications. Drought stressed plants were harvested after  
 cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d  
 and 5d). Set 3 were grown under the standard conditions  
 and then were cold stressed by placement at 4 C. Cold  
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,

and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
 2d. Set 4 were grown under the standard conditions and  
 then were heat stressed by placement at 35 C. Heat  
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
 2d and 4d and heat-stressed roots were harvested at 6 hr,  
 12 hr, 1 d, and 4d. RNA was isolated from all tissues and  
 equal RNA from each tissue and stress was pooled to  
 construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:  
 Pred. No.: 5.15e-126 length: 886  
 Score: 1260.00 Matches: 244  
 Percent Similarity: 94.29% Conservative: 20  
 Best Local Similarity: 87.14% Mismatches: 16  
 Query Match: 77.63% Indels: 0  
 DB: 14 Gaps: 0

US-09-931-457a-31 (1-325) x CK265667 (1-886)

QY 1 MetAlaValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrPro 20  
 DB 42 ATGGCGGGGAAAAGATTGGAATTGCCAAGGATGTAATGATGATTGTAACACTCTCT 101  
 QY 21 LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40  
 DB 102 TTGGTATACCTGAATATATGTTGTGATGAGGTGTGTTGCAAGTGTGCTGCAAGCTGGAA 161  
 QY 41 LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60  
 DB 162 AGCATGAGACCATGCTCTAGTGTAAAGATAGATCGGTTATGATGATTACAGATGCT 221  
 QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80  
 DB 222 GAGAGAAAGGCTCTCATCAAAACCCGGAGAGTGTCTCATCAAGCTCAAGCTGGAAC 281  
 QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleThrMet 100  
 DB 282 ACTGGTGTAGATTTGCAATTCATGCTGCTGCTAAAGCTCAAACTCATCATTCAGATG 341  
 QY 101 ProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120  
 DB 342 CTTCTCATGATGAGTCTGAGAGAAAGATTAATTCGCGTCTTCGGTGTGATGCTG 401  
 QY 121 LeuThrAspProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluGluLeuAla 140  
 DB 402 CTTACCGATCCAGCAAAAGGATGAAGGCTGCTATTCAAAAGCTGAAGAGATAAAGCC 461  
 QY 141 LysThrProAsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHis 160  
 DB 462 AAACACCCCACTCCTATATCTTCAGCAATTGAAAACCTGCTAACCCAAAGATACAC 521  
 QY 161 TyrGluThrThrGlyProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheVal 180  
 DB 522 TATGAGACCACTGCTCTGAGATCTGAAAAGCTCAATGGGAAAGTGATGCTTAGTC 581  
 QY 181 SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsn 200  
 DB 582 TCTGGAATTGAAACAGAGGACGATTAAGTGTTCAGGCAATATTGAGAGAGAGAAC 641  
 QY 201 ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyGlyLys 220  
 DB 642 CCCAAGCTTAAGCTGTATGCGGTGAACCAAGTGAAGTGTATCTTCTGTGGAAG 701  
 QY 221 ProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240  
 DB 702 CTTGTCACATTAAGATTCAGGGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 761  
 QY 241 AsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260  
 DB 762 AACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 821  
 QY 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla 280

Mon May 24 08:18:52 2004

us-09-931-457a-31.p2n.rst

Page 15

Db 822 CTGGCATTGAAGAGAGATTGCTAGTGGGAAATTCATCTGCTGCTGCGCGCGCA 881

Search completed: May 22, 2004, 00:35:05  
Job time : 3109 secs



DB 1 ATGAGAGGACCGGTTGATCAAGATGATGCAACTCAATTGATAGTAACACCCCA 60  
 QY 21 LeuValTyrLeuAsnIysLeuAlaAspGlyCysValAlaArgValAlaAlaIysLeu 40  
 DB 61 ATGATATATCTGAACACATTTGGATGTTGTAGCTGATTGACGTAAGCTTGA 120  
 QY 41 LeuMetGluProCysSerSerValIysAspArgIleGlyTyrSerMetIleAlaAspAla 60  
 DB 121 ATGATGAGCCTTGTCTAGTGTCAAGAGAGCAAGTATGCTTATGATGATTAAGATGCA 180  
 QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValIleGluProThrSerGlyAsn 80  
 DB 181 GAAGACAAAGGATGATTAATCTCTGGAGAGACACGATGAGGCTACCTCTGTAAC 240  
 QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrIysLeuIleIleThrMet 100  
 DB 241 ACCGGGATTGTTAGCTTCACTCGGTGACAGCTAAAGCTTAAAGTGTCTCTCAACATG 300  
 QY 101 ProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120  
 DB 301 CCTCATCAATGAGCCTTGAGAGAAATATCATTTTACGATTAGTGGCGAGCTTAC 360  
 QY 121 LeuThrAspProAlaIysGlyMetIysGlyAlaValGlnIysAlaGluGlnIleLeuAla 140  
 DB 361 CTCACAGATCTAGTAAGCGCTTCAAGGATATATCAACAAAGCTGAAGAGATATGTAGC 420  
 QY 141 LysThrProAsnAlaTyrIleLeuGlnIlePheGluAsnProAlaAsnProLysValHis 160  
 DB 421 AAAAATCCAGATAGATATCTGTAAGCAAGCTTCAAAAATCTTCAAAACCCGCAAACTCAT 480  
 QY 161 TyrGluThrThrGlyProGluIleTyrPlySerGlySerAspGlyLysIleAspAlaPheVal 180  
 DB 481 TATCGAACCAACGGGTCCAGAGATATGAGAGACTGTGCAAGGGAAGTACATATTGTT 540  
 QY 181 SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsn 200  
 DB 541 GCCGGTGTGGAACCTGTGGAACGCTTCCGGATCAAGAAATCTCTCAAGCAAGAAAT 600  
 QY 201 ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLys 220  
 DB 601 AAAGACTTTAAGGTTATGTGTGTGGAACCTACAGAAAGTCCGGTAATAGTGAAGCAAA 660  
 QY 221 ProGlyProHisIysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240  
 DB 661 CCGGTACACATTTGATTCAGAGTATGCGGTGACTCATCCAGACCAATTGATTTTC 720  
 QY 241 AsnLeuLeuAspGluValAlaGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260  
 DB 721 AACCTTCTGATGATGATCAAGTATGAGTGAACCAAGTGAAGCAATTGAACAGCAAACTT 780  
 QY 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla 280  
 DB 781 CTGCCCCGAAGAGATTAATCTGCTGAGGATATCTCTGAGCTGTGCGAGCGCTGCG 840  
 QY 281 PheGlnIleAlaIysArgProGluAsnAlaGlyLysLeuIleValAlaAlaPheProSer 300  
 DB 841 ATAAAGGTGCAAGCGCCAGAAACGCCGCAAACTCATAGTTGTGATTTTCCCTAGC 900  
 QY 301 PheGlyLysGlyTyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSer 320  
 DB 901 GGTGGGGAAGCTTACTTATGACCTCATTTGTAATCACTGACATGAGACATGAGACGAAT 960  
 QY 321 MetThrPheGlu 324  
 DB 961 TTGCCAATTCAA 972  
 RESULT 2  
 US-09-103-840A-2  
 ; Sequence 2, Application US/09103840A  
 ; Patent No. 6294328  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FLEISCHMAN, Robert D.  
 ; APPLICANT: WHITE, Owen R.

; APPLICANT: FRASER, Claire M.  
 ; APPLICANT: VENTER, John C.  
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
 ; TITLE OF INVENTION: TUBERCULOSIS  
 ; FILE REFERENCE: 24366-20007.00  
 ; CURRENT APPLICATION NUMBER: US/09/103,840A  
 ; CURRENT FILING DATE: 1998-06-24  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 4403765  
 ; TYPE: DNA  
 ; ORGANISM: Mycobacterium tuberculosis  
 ; FEATURE:  
 ; OTHER INFORMATION: CDC 1551  
 ; OTHER INFORMATION: "n" bases at various positions throughout the sequence  
 ; OTHER INFORMATION: represent a, t, c or g  
 US-09-103-840A-2  
 Alignment Scores:  
 Pred. No.: 7,71e-97 Length: 4403765  
 Score: 949.50 Matches: 181  
 Percent Similarity: 76.95% Conservative: 56  
 Best Local Similarity: 58.77% Mismatches: 68  
 Query Match: 58.50% Indels: 3  
 DB: 3 Gaps: 3  
 US-09-931-457A-31 (1-325) x US-09-103-840A-2 (1-4403765)  
 QY 8 IleAlaIysAspValThrGluLeuIleGlyLysThrProLeuValTyrLeuAsnLysLeu 27  
 DB 2604610 ATCGCCGAGACATCAACCACTATCGGGCGCACACCGCTGGCTGCACTGCGCGAGATC 2604669  
 QY 28 AlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluProCysSerSer 47  
 DB 2604670 ACCGACGGCGCGTTCGCGACATCTGCGCAAGCTGAATTCTTCAACCGCGCAACAGC 2604729  
 QY 48 ValIysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGlnLysGlyLeuIleThr 67  
 DB 2604730 GTAAAGACCGTATCGGGTTCATGCTCCAGCGCGGACGACGAGCTTTGATCAAG 2604789  
 QY 68 ProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIleGlyLeuAlaPhe 87  
 DB 2604790 CCG--GACACGATCATCTCGAACCCACGAGCGGTAAACCGGCATCGCCCTGGCATG 2604846  
 QY 88 MetAlaAlaAlaArgGlyTyrIysLeuIleIleThrMetProAlaSerMetSerLeuGlu 107  
 DB 2604847 GTTTCGCGGCGACCGGCTACCGGTGCTGCTGACCATGCGCGAGACGATGAGTGTGAG 2604906  
 QY 108 ArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspProAlaLysGly 127  
 DB 2604907 CGCGGATGTGCTGCGCGCATACGCTGTAATCATCTCACTCCGGGTGCGAGCGGC 2604966  
 QY 128 MetIysGlyAlaValGlnIysAlaGluIleLeuAlaLysThrProAsnAlaTyrIle 147  
 DB 2604967 ATGTCAAGTGCATCGCCAAAGCTGAGAG--CTGGCCAAAGACCGATCAACGCTACTTC 2605023  
 QY 148 Leu--GlnIlePheGluAsnProAlaAsnProLysValHisTyrGluThrThrGlyPro 166  
 DB 2605024 GTGCCCCGCAATTCAGAGACCGCGGAACCGGCATTCATCGCTCAACGACCGCGAGC 2605083  
 QY 167 GluIleTyrPlySerGlySerAspGlyLysIleAspAlaPheValSerGlyIleGlyThrGly 186  
 DB 2605084 GAGGTCTGCGGTGACACCGACGCAAGTGCATGCTGCTGCGGAGTGGACCGGT 2605143  
 QY 187 GlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsnProAsnIleLysLeuIle 206  
 DB 2605144 GGCACCATCAACCGGCTGCGGACGTCATCAAGAGACGCAACCGCTCGCGCGGTTCGTG 2605203  
 QY 207 GlyValGluProValGluSerProValLeuSerGlyLysProGlyProHisIysIle 226  
 DB 2605204 GCCGTAGACCGCGCGCGCTGCTCTTCTGTGTGCTCAGAGAGGACCGCACCCGATC 2605263



[illegible]

### RESULT 3

```

; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

```

Alignment Scores:

Pred. No.:	7 74e-97	length:	441522
Score:	949.50	Matches:	181
Percent Similarity:	76.95%	Conservative:	56
Best Local Similarity:	58.77%	Mismatches:	68
Query Match:	58.50%	Indels:	3
DB:	3	Gaps:	3

US-09-931-457A-31 (1-325) x US-09-103-840A-1 (1-4411529)

[illegible]

QY 108 ArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspProAlaTyrGly 127  
 Db 2609097 CGCCGATGTTGCTCGCGGCATACGGTCTGAATCATCTCACTCCGGTGCCGAGCGC 2609156  
 QY 128 MetIleGlyAlaValGlnLysAlaGluIleLeuAlaLysThrProAsnAlaTyrIle 147  
 Db 2609157 ATGTCAAGTGCATCGCCAAAGCTGAGAG--CTGSCCAAGACCGATCAACGCTACTTC 2609213  
 QY 148 Leu--GlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThrGlyPro 166  
 Db 2609214 GTGCCCCAGCAATTCAGAAACCCGGCGAACCCGGCCATCCATCCGCTCAAGACCGCCGAG 2609273  
 QY 167 GluIleTrrPlySglySerAspGlyLysIleAspAlaPheValSerGlyIleGlyThrGly 186  
 Db 2609274 GAGGTCTGGCTGACACCAACGACGGCAAGSTCGACATCGTCTGCGCGGAGTCGGCAACCGGT 2609333  
 QY 187 GlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsnProAsnIleLysLeuIle 206  
 Db 2609334 GGCACCATCAACCGCGCTGCGCAGGTCAATCAAGAAACGCAACGCTCGGCCCGTTCTGTG 2609393  
 QY 207 GlyValGluProValGluSerProValLeuSerGlyGlyLysProGlyProHisLysIle 226  
 Db 2609394 GCCGTAGAGCCGACCGCGCTGCGCGGCTCTTCTGTGTGGCCAGAGGGAACCGCAACCCGATC 2609453  
 QY 227 GlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeuLeuAspGluVal 246  
 Db 2609454 CAGGCGATCGCGCGCGGGTTCGTCCCGCGGTACTCGAACCGAACCTTAGTCGACGAGATC 2609513  
 QY 247 ValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeuLeuAlaLeuLysGluGly 266  
 Db 2609514 ATTACCGTCCGTAAACGAAGACGCGCTCAACGTGGCGCGCGCGCTGGCCCGGGAAGAGGCG 2609573  
 QY 267 LeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAlaAlaPheGlnIleAlaLysArg 286  
 Db 2609574 TTGCTGTGTGCGCATCTCCTCGGCGCGCCCAAGTGGCCGCTCTTCAAGTGGCCCCGCGCG 2609633  
 QY 287 ProGluAsnAlaGlyLysLeuIleValAlaValaPheProSerPheGlyGluArgTyrLeu 306  
 Db 2609634 CCAGAGAAACGCCGGAAGACTTAATCGTGTAGTGTCTCCCGACTTCGCGGAACGATATCTG 2609693  
 QY 307 SerSerValLeuPheGluSerVal 314  
 Db 2609694 AGCACACCGCTTCTCGCCGACGTG 2609717

## RESULT 4

```

US-09-596-002-39
; Sequence 39, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, B.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596, 002
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 39
; LENGTH: 100848
; TYPE: DNA
; ORGANISM: M. catarrhalis
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte template ID No. 6632636 39
; PUBLICATION INFORMATION:
US-09-596-002-39

```

**Alignment Scores:**

Pred. No.:	1.29e-95	Length:	100848
Score:	917.00	Matches:	182

Mon May 24 08:18:51 2004

us-09-931-457a-31.p2n.mri

Page 4

Percent Similarity:	76.72%	Conservative:	52
Best Local Similarity:	59.67%	Mismatches:	67
Query Match:	56.50%	Indels:	4
DB:	4	Gaps:	3

US-09-931-457A-31 (1-325) x US-09-596-002-39 (1-100848)

QY	8	IIAlAlAlAspValThrGluLeuIleGlyLysThrProLeuValTyrLeuAsnLysLeu	27
DB	17953	ATTGCAAAAAGTGTGACTGTACCTTATCGGTAAACCCGCTGTGTCAACTCAATCGCCTC	18012
QY	28	AlAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluProCysSerSer	47
DB	18013	ACCGAAGCCTTATCTGCTGCTGTGTGTCAAGCTTGAGTATGCCAATCCAGCAAGTTCG	18072
QY	48	ValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGluLysGlyLeuIleThr	67
DB	18073	GTCAAGGATCGCATTTGCCCTTCTTATGATGACGATGCCGAAAATCAGGGCTTTTGACC	18132
QY	68	ProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIleGlyLeuAlaPhe	87
DB	18133	AA--GACACCACCATCGTTGAGCGCAGCAGTGGCAATACAGGCATTGGGCTTGCCATG	18189
QY	88	MetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProAlaSerMetSerLeuGlu	107
DB	18190	GTGGCAGCGCGCAGAGGTACCGCTGTGTATTCATGCTGAGAGCATGAGCCTTGAG	18249
QY	108	ArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspProAlaLysGly	127
DB	18250	CGTCGTGCTGTGCTGTGTCGATATGGGCCAAGCTGTGTGACGCCAGCGGTGAGGCGC	18309
QY	128	MetLysGlyAlaValGluLysAlaGluGluIleLeuAlaLysThrProAsnAlaTyrIle	147
DB	18310	ATGGGTGGCGCATGTCCAAGCCAAATGAGCTGCTGCCCAA----GACGGCTATTTT	18363
QY	148	Leu--GlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThrGlyPro	166
DB	18364	ATGCCACGCCCAATTGTATTAATTAGCCATTCACAATCCACCGTACACAACAAGCCGAA	18423
QY	167	GluIleTyrPheGlySerAspGlyLysIleAspAlaPheValSerGlyIleGlyThrGly	186
DB	18424	GAGATTGGCGCGATACAGATGTCAAGTGTATTTGTCTCAGGTGTGGTACAGGT	18483
QY	187	GlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsnProAsnIleLysLeuIle	206
DB	18484	GGTACAGTGCAGGTGTGCGGTGAGACACTCAAGCCAAAAGAGCCATTCAAAATTGTT	18543
QY	207	GlyValGluProValGluSerProValLeuSerGlyLysProGlyProHisLysIle	226
DB	18544	GCAAGTTCAGCCATTGGCATCACTGTCTTATCAGCGGAGAAAAAGGCCGACCCCAATT	18603
QY	227	GlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeuLeuAspGluVal	246
DB	18604	CAAGGCATTGGTGCAGGGTTTATCCCATCAGTACTCAATACGATATCATGATGAGTTC	18663
QY	247	ValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeuLeuAlaLeuLysGluGly	266
DB	18664	ATCAAGTCGCTTAATGAAGATGCCCTTAGTGTGCCACGAGATATGGCACCACAAGAGGC	18723
QY	267	LeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAlaAlaPheGlnIleAlaLysArg	286
DB	18724	TTGCTGTGTGTATCTCATCAGGTGCGGCGAGTTTCGGCGGCACTGCAATTGGCGGACGA	18783
QY	287	ProGluAsnAlaGlyLysLeuIleValAlaValPheProSerPheGlyGluArgTyrLeu	306
DB	18784	GAAGAAAACGCGAGTAACTTATCGTGTATCATACCATCATCAGGTGACGCTTATTG	18843
QY	307	SerSerValLeuPhe 311	
DB	18844	TCAACAGCATTAATT 18858	

```

: Sequence 3561, Application US/09543681A
: Patent No. 6605709
:
: GENERAL INFORMATION:
: APPLICANT: GARY BRETON
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
: TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 2709.1002-001
: CURRENT APPLICATION NUMBER: US/09/543,681A
: CURRENT FILING DATE: 2000-04-05
: PRIOR APPLICATION NUMBER: US 60/1128,706
: PRIOR FILING DATE: 1999-04-09
: NUMBER OF SEQ ID NOS: 8344
: SEQ ID NO 3561
:
: LENGTH: 1008
:
: TYPE: DNA
:
: ORGANISM: Proteus mirabilis
:
: US-09-543-681A-3561

```

Alignment Scores:	1.13e-94	Length:	1008
Pred. No.:	882.00	Matches:	181
Score:	72.50%	Conservative:	51
Percent Similarity:	56.56%	Mismatches:	76
Best Local Similarity:	54.34%	Indels:	12
Query Match:	4	Gaps:	5
DB:			

US-09-931-457A-31 (1-325) X US-09-543-681A-3561 (1-1008)

QY	3	ValGluArgSerGlyIleAlaIysAspValThrGluLeuIleGlyLysThrProLeuVal	22
Db	49	ATAAAATGAGCAAAATTTTCGAGATTAACTCGCAAACTATTGGCCATACCCCTAGTT	108
QY	23	TyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGluMet	42
Db	109	CGTTTAAACATTTCCGTATGCG-----AATATCTTAGCTAAAGTAGAATCTCGT	159
QY	43	GluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGlu	62
Db	160	AATCCAAGCTTTAGTGTAAATGCCGTATCGGTGCAAAATATGATTGGGATGCAGAAAAA	219
QY	63	LysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGly	82
Db	220	AAAGTATCTCTGAACCCGCGCTAGAG---CTAGTTAGCCCAACCACTGGTATATACTGGG	276
QY	83	IleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProAla	102
Db	277	ATTGCCCTTGCTTATGTAGCTGCTGCGCGCTGCTATAGCTGACACTGACTATGCTCGT	336
QY	103	SerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThr	122
Db	337	ACCATGAGTATTGAGCGCGGTAAATTATTAAAGCAATTAGGTGCTAACCTTGTGTTAACT	396
QY	123	AspProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluGluIleLeuAlaLysThr	142
Db	397	GAGGGCGCTAAAGGTATGAAGGAGCGATTGATTAAGCGAAGCAATTCGTGATAGCGAT	456
QY	143	ProAsnAlaTyr---IleLeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyr	161
Db	457	CCTAGCCGTTACCTTCTACTACACAATTTAGTATCCAGCAAAACCCAGAAATTCATGAG	516
QY	162	GluThrThrGlyProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheValSer	181
Db	517	AAACACACAGGCCCTGAATTTGCAATGACACTGATGGCAAAAGTTGATGTAGTTATGCA	576
QY	182	GlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGlu---GlnAsn	200
Db	577	GGTGTCCGAACGGGTGTACCATCAACGGGTATTGTCGTATTGAAAAATACTCAAGGT	636
QY	201	ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyGly---	219
Db	637	AAAGCTGTCAACCATGGTTGGGTAGAACCTTAAAGATTCTCTGTTATTAGCCACGACATTA	696
QY	220	-----LysProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIle	234

Db 697 GCAAGTGAAGAATTAAGCTGGTCCACATAAATCCAGGTATGTGACAGTTTATC 756  
Qy 235 ProGlyValLeuGluValAsnLeuLeuAspGluValValGlnIleSerSerAspGluAla 254  
Db 757 CCTGAAGAACTTAGAGTTATCACTATTAGATCGCGTGAATCCAAATTACTAACAAGAGAACA 816  
Qy 255 IleGluThrAlaLysLeuLeuAlaLeuLysGluGluLysPheValGlyIleSerSerGly 274  
Db 817 TTGAAACAGCTCGTACCGTTATGACCAAGAAGGTATTTAGCGGTATTCATCCGT 876  
Qy 275 AlaAlaAlaAlaAlaAlaPheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIle 294  
Db 877 GCGGCAATTGCGGCTGCAGTGAATTAAGCTTAAGAGCCAGAAATTGCTAATAAGGAATC 936  
Qy 295 ValAlaValPheProSerPheGlyGluArgTyrLeuSerSerValLeuPheGluSerVal 314  
Db 937 GTAGTGAATTTTACCTTCTCAGGTGAACGTATCTCAGCACCCCACTTTTGTGATATC 996

## RESULT 6

US-09-328-352-1324  
; Sequence 1324, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNI FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 1324  
; LENGTH: 1008  
; TYPE: DNA  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-1324

Alignment Scores:  
Pred. No.: 1.16e-92 Length: 1008  
Score: 865.00 Matches: 179  
Percent Similarity: 72.64% Conservative: 52  
Best Local Similarity: 56.29% Mismatches: 75  
Query Match: 53.30% Indels: 12  
Gaps: 6

US-09-931-457A-31 (1-325) x US-09-328-352-1324 (1-1008)

Qy 7 GlyIleAla-----LysAspValThrGluLeuIleGlyLysThrProLeuValTyrLeu 24  
Db 49 GGTATCGCAGTTTATTCAATTAATGACAGAGCTATTGTGAATACCTCTTAGTAGCTATT 108  
Qy 25 AsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluPro 44  
Db 109 AATCGTTGATTAA--ACTGATGCACTGTATTAGCCAAAGTAGAAAGTCGTACCCA 165  
Qy 45 CysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGluLysGly 64  
Db 166 GCTTTTCTGTGAATGCCGTAATGAGCTCGCTGATTGCAATGCAAGAAAAAGAGGC 225  
Qy 65 LeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIleGly 84  
Db 226 GCGTTGAAGAGGAGCATAT--GTGAACCAACCAAGGTATTAATCTGATTTGCT 282  
Qy 85 LeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProAlaSerMet 104  
Db 283 TTGGCGTTGTCGAGCGCGGCAAAAGTTATTCATTACTTAATGCTGCTGCAAGCATG 342  
Qy 105 SerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspPro 124  
Db 343 AGCCTTGAACGTAGAAAAGTGTAAAGCACTCGTGCCAATTAGTCTTAACAGAACCA 402  
Qy 125 AlaLysGlyMetLysGlyAlaValGlnLysAlaGluGluIleLeuAlaLysThrProAsn 144

Db 403 GCGAAGGAATGAAGCTGCAGTTGATGAAGCTGTACGATTAGCAACTGACAGCAGAA 462  
Qy 145 AlaTyrIleLeu----GlnGlnPheGluAsnProAlaAsnProLysValIleTyrGluThr 163  
Db 463 GTTATTCTTCCCAACAAGTTGAAAACCACTATCCACAATTCATGTTGATACG 522  
Qy 164 ThrGlyProGluIleThrLysGlySerAspGlyLysIleAspAlaPheValSerGlyIle 183  
Db 523 ACTGACCAAGAAATCTGCGCAAGCAACAGGTGGCAAGTTGATATCTTGTGCTGCTGTA 582  
Qy 184 GlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGlu---GlnAsnProAsn 202  
Db 583 GGTACAGAGAGTACGATTAACGGGTATTTCTCGTTATTTGAACAAGTTCAAGATAAGCCA 642  
Qy 203 IleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLys----- 220  
Db 643 TTGTACTCAGTAGCAGTTGAACCTGCTGAGCTTCCAAATTATTAATCAACAAAAATGCT 702  
Qy 221 -----ProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGly 236  
Db 703 GAGAACATTACACCCGACCGCATAAATTCAGGAATTCAGCAAAATTTATTCCAAAG 762  
Qy 237 ValLeuGluValAsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGlu 256  
Db 763 AATCTGATTAGATTGGTTGATGAAGTCTTAACAGTAGAGTAGAAGCAATTCAA 822  
Qy 257 ThrAlaLysLeuLeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAla 276  
Db 823 TGGCGAAGAAAGTGTGCCAACCAAGAGTATTAGTGGTATTCAAGTGGCTGCA 882  
Qy 277 AlaAlaAlaAlaPheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAla 296  
Db 883 TTGCACAGAGTGGCTAAATTCAGAACGCTCTGAATAATGACAGTAACAATTTGTTGT 942  
Qy 297 ValPheProSerPheGlyGluArgTyrLeuSerSerValLeuPheGluSerVal 314  
Db 943 ATTCTTCAGATAGTGGGAACGCTATTATCTTCTGTTTATTGAAGTTTA 996

## RESULT 7

US-09-252-991A-14920  
; Sequence 14920, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 14920  
; LENGTH: 2247  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-14920

Alignment Scores:  
Pred. No.: 4.83e-91 Length: 2247  
Score: 856.00 Matches: 182  
Percent Similarity: 70.81% Conservative: 46  
Best Local Similarity: 56.52% Mismatches: 84  
Query Match: 52.74% Indels: 10  
Gaps: 6

US-09-931-457A-31 (1-325) x US-09-252-991A-14920 (1-2247)

Qy 6 SerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrProLeuValTyrLeuAsn 25  
Db 1276 AGCCGATCTTCGACAGCAACGCGCAGTCCATGCGCAACACGCGCTGTTCCAGATCAAC 1335



```
QY 26 LysLeuAlaaspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluProCys 45
DB 1336 CGCATCGCCCGCGCGCGCTCACC--ATCTCGGCCAAGATCGAAGGGCGCAACCCGCGC 1392
QY 46 SerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGluLysGlyLeu 65
DB 1393 TATTCGGTGAAGTCCGGATCGCGCGCAACATGATCTGGATGCGCAAGCCAGCGGCAGA 1452
QY 66 IleThrProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIleGlyLeu 85
DB 1453 CTGAAGTCCGGCATGACC--CTGCTCGAGCCGACCTCCGGCAACACCGGCATCGGCCTC 1509
QY 86 AlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProAlaSerMetSer 105
DB 1510 GCCTTCGTCGCCCGCGCAACGCGCTCACTCAACTGATCTGACCATGCGCGCGTGTGATGAC 1569
QY 106 LeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspProAla 125
DB 1570 CTGAGCGCAGCAGAGGTCTCAAGCGCTGGCGCGCGAGCTGTCTCTGACCGAACCAGCGCG 1629
QY 126 LysGlyMetLysGlyAlaValAlaGlnLysAlaGluGluIleLeuAlaLysThrProAsnAla 145
DB 1630 AAGGCGATGAAGGTCTCGATCCAGAGGCGCGCAACTGCTGCGCGGGGATCCCGGGAAG 1689
QY 146 TyrIleLeu--GlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThr 164
DB 1690 TACTTCATGCCGCGCAGCTTCGACCAACCCGCGCAACCCGCGCATCTGAGGAAGACCAACC 1749
QY 165 GlyProGluIleTrpLysGlySerAspGlyLysIleAspAlaPheValSerGlyIleGly 184
DB 1750 GGGCGGAGATCTGGAACGATACCGAGGCGCGGTGACGTGCTGTCTCCGGGGTCTGGC 1809
QY 185 ThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGlu--GlnAsnProAsnIle 203
DB 1810 ACCGCGGACCTCACCCTGCTCTGCGCTACATCAAGAACACCCGGGCAAGCCGATC 1869
QY 204 LysLeuIleGlyValGluProValGluSerProValLeuSer-----GlyGly--- 219
DB 1870 CTCGCCGTGAGGCTGAGCGCTGACCTCGCGGTGATCAGCCAGACCTCGCGCGGAG 1929
QY 220 -----LysProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyVal 237
DB 1930 GAGGTCAAGCCCGCGCGCAAGATCCAGGCGCATCGCGCGCTTCTGTCGGAAGAAC 1989
QY 238 LeuGluValAsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThr 257
DB 1990 CTCGACCTGTCTGCTGATCGACCGGCTCGAGAAATCGGCGACGAGCGAAGAACATG 2049
QY 258 AlaLysLeuLeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAla 277
DB 2050 GCGCTCGCCTGATGACGAGAAGGATCTCTGCGGCATCTCTCGCGCGCGCGGATG 2109
QY 278 AlaAlaAlaPheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaVal 297
DB 2110 GCGCGCGCGGTGCGCTGCGCGCGAGAGCGCAACATCGAGCGCAAGACCATGCTGTATC 2169
QY 298 PheProSerPheGlyGluArgTyrLeuSerSerValLeuPheGluSerValArgGlu 317
DB 2170 CTCCTCCGATTCGCGCGCGCTATCTTTCGAGCATGCTCTTCGACGCGCTTCAGCGAA 2229
QY 318 AlaGlu 319
DB 2230 CAGGAA 2235
RESULT 8
US-09-252-991A-14450/c
; Sequence 14450, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
```

```
FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14450
; LENGTH: 2334
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14450
Alignment Scores:
Pred. No.: 5.14e-91 Length: 2334
Score: 856.00 Matches: 182
Percent Similarity: 70.81% Conservative: 46
Best Local Similarity: 56.52% Mismatches: 84
Query Match: 52.74% Indels: 10
DB: Gaps: 6
US-09-931-457A-31 (1-325) x US-09-252-991A-14450 (1-2334)
QY 6 SerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrProLeuValTyrLeuAsn 25
DB 1125 AGCCGATCTTCGAGACACGCGCAGTCATCGGACACGCGCTGTCCAGATCAAC 1066
QY 26 LysLeuAlaaspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluProCys 45
DB 1065 CGCATCGCCCGCGCGCGTCAACC--ATCTGCGCAAGATCGAAGGGCGCAACCCGCGC 1009
QY 46 SerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGluLysGlyLeu 65
DB 1008 TATTCGGTGAAGTCCGATCGCGCGCAACATGATCTGGATGCGCAAGCCAGCGGCAGA 949
QY 66 IleThrProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIleGlyLeu 85
DB 948 CTGAAGTCCGGCATGACC--CTGCTGAGCGGACCTCCGCAACACCGGCATCGGCCTC 892
QY 86 AlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProAlaSerMetSer 105
DB 891 GCCTTCGTCGCCCGCGCGCAAGCTCAACAATGATCTGACCATGCGCGCGTGTGATGAC 832
QY 106 LeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspProAla 125
DB 831 CTGAGCGACGCAAGGTCTCAAGCGCTGCGCGGAGCTGTCTGACCGAACCAGCGC 772
QY 126 LysGlyMetLysGlyAlaValAlaGlnLysAlaGluGluIleLeuAlaLysThrProAsnAla 145
DB 771 AAGGCGATGAAGGTGCGATCGAGAGCGCGAAGAACTGTGCGCGGATCCCGGAAG 712
QY 146 TyrIleLeu--GlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThr 164
DB 711 TACTTCATCGCGCAGAGTTCGACCAACCGCGCAACCGCGCATCGAGAGACCAACC 652
QY 165 GlyProGluIleTrpLysGlySerAspGlyLysIleAspAlaPheValSerGlyIleGly 184
DB 651 GCGCGGAGATCTGGAACGATACGAGGCGCGGTGAGCGTGTCTGCGGGGTCTGCGG 592
QY 185 ThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGlu--GlnAsnProAsnIle 203
DB 591 ACCGCGGACCTCACCCTGCTCGCGCTACATCAAGAACACCCGGGCAAGCGCATC 532
QY 204 LysLeuIleGlyValGluProValGluSerProValLeuSer-----GlyGly--- 219
DB 531 CTCGCCGTGAGCGGTGAGCGCGGTGACCTCGCGGTGATCAGCCAGACCTCGCGCGAG 472
QY 220 -----LysProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyVal 237
DB 471 GAGGTCAAGCCCGCGCGCAAGATCCAGGCGCATCGCGCGCGCTTCTGTCGGAAGAAC 412
QY 238 LeuGluValAsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThr 257
```



Db	411	CTCGACCTGTGCTGTGCTGACCGGGTTCGAGAAGATCGGCGACGACGAAGCGAAGACATG	3522
QY	258	AlAlysLeuLeuAlaLeuLySGluGlyLeuPheValGlyIleSerSerGlyAlaAla	277
Db	351	GGCGTGGCGCTGATGCAGAGAAGGCATCTCTGCGGCATCTCCTCCGCGCGCGCATG	2922
QY	278	AlAlaAlaPheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaVal	297
Db	291	GCGCGCGCGGTGCGCGCTGCGCCGAGGACCGCAATGCAGGCGAAGACCATCGTGTGATC	232
QY	298	PheProSerPheGlyGluArgTyrIleuSerSerValIleuPheGluSerValArgArgGlu	317
Db	231	CTCCCGCATTCGGCGGAGCGCTATCTTCAGACATCTCTTCGACGCGCTGTTCAGCGAA	172
QY	318	AlaGln 319	
Db	171	CAGGAA 166	

```

      RESULT 9
      US-09-489-039A-1995
      ; Sequence 1995, Application US/09489039A
      ; Patent No. 6610836
      ; GENERAL INFORMATION:
      ; APPLICANT: Gary Breton et. al
      ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
      ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
      ; FILE REFERENCE: 2709.2004001
      ; CURRENT APPLICATION NUMBER: US/09/489,039A
      ; CURRENT FILING DATE: 2000-01-27
      ; PRIOR APPLICATION NUMBER: US 60/117,747
      ; PRIOR FILING DATE: 1999-01-29
      ; NUMBER OF SEQ ID NOS: 14342
      ; SEQ ID NO 1995
      ; LENGTH: 981
      ; TYPE: DNA
      ; ORGANISM: Klebsiella pneumoniae
      US-09-489-039A-1995
    
```

Alignment scores:	
Pred. No.:	2,54e-91
Score:	853.50
Percent Similarity:	71.52%
Best Local Similarity:	54.18%
Query Match:	52.59%
DB:	4
	Gaps: 5
	Length: 981
	Matches: 175
	Conservative: 56
	Mismatches: 79
	Indels: 13
	Gaps: 5

US-09-931-457A-31 (1-325) X US-09-489-039A-1995 (1-981)

OY		6	SergIylIEalAlalsaspvalThrgluleuilegllylsThrProleuValIyrleuAsn	25
Db		13	AGTAAATTTTGAAGACAACCTCCCTGACTATCGGTGCATACACC GGCTGGTAGCACTGAAC	72
OY		26	LysleuAlaspglycysValAlaargValalaAlalsleuGluleumetGIuProcys	45
Db		73	CCCATCGTAATGCC-----CGTATCCTGGCAAAGGTAGAATCTCGTAACCCGAGC	123
OY		46	SerSerVallysaspargileglyTyrsertmetleAlaspAlagluGIuLysglyLeu	65
Db		124	TTTAGCGTAATAATGCCGTATCGCGCCAATATGATCTGGAGTGCTGAAAAGCGCGCGTT	183
OY		66	IleThrProGIlylYsserValleuileGIuProthrSergIyasnThrglyleGIuLeu	85
Db		184	CTGAACCTGGCGGTGAA--CTGGTCGAGCGCACCAAGCGCAACACCGTATCGCCCTC	240
OY		86	AlaphemetalAlaAlaIaargglyTyryrIsleuilelleThrm etProAlasermetSer	105
Db		241	GCTAACGTGGCGGGGGCCCGCGCTACAACGTAGACGCTAACCATGCGGAACCATGAGC	300
OY		106	LeuGIuArgArgillelleleuLeuAlaphedilyAlagluLeuValleuthrasPProAla	125
Db		301	ATCGAAGCGTCGTAGCTGTGAAAACCTGGGGGCCCAACCTGTGTCTGACCGAAGGGCC	360

OY		126	LyeglymettlysGlyAlaValGlnLysAlaGluGluLeuAlaIleThrProAsnAla	145
Db		361	AAAGGCATGAAGGCCCTATCCAGAAGCCGAAGAGATTGTCCGCACCAACC GGAAACA	420
OY		146	Tyr--IleLeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThr	164
Db		421	TTTCCTGCTGCTGACAGCAGTTTCAGCAACCCGGCTAACCCGGAAATCCACGAGAATACTACC	480
OY		165	GlyProGluIleTryptylseraspGlyLysIleaspAlaPheValSerGlyIleGly	184
Db		481	GGCCCGGAAATCTGGGAAGATACCGATGCTCAGGTGATGTGTTATCTCCGGCGTAGCT	540
OY		185	ThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsnProAsnIleLys	204
Db		541	ACCGCGGAACGCTGACCGCGCTCACCGCTATATCAAAAACACCAAGGTAAAAAAGAT	600
OY		205	IleIle-----GlyValGluProValGluSerProValLeuSerGlyGly-----	219
Db		601	CTGATCACCGTCGCGCGTTGAACCGACCGATTCCCGGTGATGCGCAGCGCTGGCGGG	660
OY		220	-----LysProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGly	236
Db		661	GAAGAGCTCAAGCCGGGTCCGCACAATACTCAGGTATCGCGCGGCTTCATTCTCTGGC	720
OY		237	ValLeuGluValAsnLeuLeuAspGluValValGlnIleSerSeraspGluAlaIleGlu	256
Db		721	AACCTGATCTGAAGCTGTGTGCACAAGTGATCGCATTAACCAACGAGAACTATCTCC	780
OY		257	ThrAlaLysLeuLeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAla	276
Db		781	ACGGCGCGCGCCCTGATGAGAGAAAGAGCATCTCTGGCGGTATCTTCCGGGGCCGCC	840
OY		277	AlaAlaAlaAlaPheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAla	296
Db		841	GTTGCCCGCGCGCTGAAGCTGCAGGAAGATGAGGCCTTTACCAATAGATAATTGTGTT	900
OY		297	ValPheProSerPheGlyGluArgTyrLeuSerSerValLeuPheGluSerValArgArg	316
Db		901	ATCTGCGCTTCTCTCAGGTGAGCGTTATTAAAGACCGCGCTGTTCGCCGATCTCTCACT	960
OY		317	GluAlaGlu 319	
Db		961	GAGAAAGAG 969	

RESULT 10  
US-09-557-884-1

; Sequence 1, Application US/09557884  
; Patent No. 6506581  
ATTORNEY, FIRM, AND ADDRESS

GENERAL INFORMATION:

APPLICANT: Fleischmann et al.  
TITLE OF INVENTION: The Nucleotide sequence of

the Haemophilus influenzae  
Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Science  
STREET: 9410 Key West Avenue

CITY: ROCKVILLE  
STATE: MD

ZIP: 20850

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell Pentium  
OPERATING SYSTEM: MS DOS v6.22  
CONTAINER: 3000T Part

```

; SOFTWARE: ASCII text
;
CURRENT APPLICATION DATA:
ADDITIONAL NUMBER: 116/06/557 884

```

APPLICATION NUMBER: 08/05/25/1,000  
FILING DATE: 25-Apr-2000  
ORIGINATOR: FBI-NEW

CLASSIFICATION: <UNKNOWN>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/435 103

APPLICATION NUMBER: 08/410,102  
FILING DATE: JUN-5-1995

ATTORNEY/AGENT INFORMATION:  
 NAME: Michelle S. Marks  
 REGISTRATION NUMBER: 41,971  
 REFERENCE/DOCKET NUMBER: PB186P3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 301-309-8504  
 TELEFAX: 301-309-8439  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1830121 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-09-557-884-1

Alignment Scores:  
 Pred. No.: 1.03e-81 Length: 1830121  
 Score: 816.50 Matches: 172  
 Percent Similarity: 70.16% Conservative: 49  
 Best Local Similarity: 54.60% Mismatches: 83  
 Query Match: 50.31% Indels: 11  
 Gaps: 6

US-09-931-457A-31 (1-325) x US-09-557-884-1 (1-1830121)

QY 8 ileAlaLysAspValThrGluLeuIleGlyLysThrProLeuValTyrLeuAsnLysLeu 27  
 DB 1165136 ATTATGACAGACATCTTATCTATCGGAATACGCCGCTTGCGTTAAACACTTT 1165195  
 QY 28 AlaSpGlyCysValAlaArgValAlaAlaLysLeuLeuMetGluProCysSerSer 47  
 DB 1165196 -----GCCCATACCGTATATGTGTGTAATAATGAAGTGTGAACCAAGCTACAGC 1165249  
 QY 48 ValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGluLysGlyLeuIleThr 67  
 DB 1165250 GTAAATGCGCTATGCGGCGAATATGTGTGCAAGCAAGAAAGATGCGACGCTCACA 1165309  
 QY 68 ProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIleGlyLeuAlaPhe 87  
 DB 1165310 AAAGGGAAGAGATT---GTAGATGCAACGAGTGTGAACACGCGCATTCCTTGGCTTAT 1165366  
 QY 88 MetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProAlaSerMetSerLeuGlu 107  
 DB 1165367 GTTCCGCGGCTAGAGGTTATTAATCAACGCTCACTATGCCAGAAACATGACTTGAA 1165426  
 QY 108 ArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspProAlaLysGly 127  
 DB 1165427 AGAAAACGCTTATTTGCGCATTTGGGTGTAATTTAGTCTTACCGAAGCGCCAAAAGA 1165486  
 QY 128 MetLysGlyAlaValGlnLysAlaGluGluIleLeuAlaLysThrProAsnAlaTyrIle 147  
 DB 1165487 ATGAAAGGTCTATTGCGAAGCAGAAAGAAATGTTGCTTCTGATCCAGCGCTATGTC 1165546  
 QY 148 ---LeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThrGlyPro 166  
 DB 1165547 ATGCTTAAACAATTGAAATTCAGCCCAACCCACAAATTCAATCGAAGAACACAGGCT 1165606  
 QY 167 GlnIleTyrLysGlySerAspGlyLysIleAspAlaPheValSerGlyIleGlyThrGly 186  
 DB 1165607 GAAATTTGAAAGATACGAGATGCGCAAGTCGATGTTGTTGCTGCGCTAGCAACAGGT 1165666  
 QY 187 GlyThrIleThrGlyAlaGlyLysTyrLeuLys---GluGlnAsnProAsnIleLysLeu 205  
 DB 1165667 GGTTCGATTACGGGCAATTCCTCGCGCATTAATTAATTTGTTGTAACAAATCACTTCT 1165726  
 QY 206 IleGlyValGluProValGluSerProValLeuSer-----GlyGly----- 219  
 DB 1165727 GTTCCGCTTGAAGCAGTGAATCTCCGCTCATTAGTCAAACTTTAGCGGGTGAAGAGTA 1165786  
 QY 220 LysProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGlu 239  
 DB 1165787 AAACCAAGCCCAACAAATTCAGAGTATCGTGGGCTTCATTTCCCAAAATTTAGAT 1165846

QY 240 ValAsnLeuLeuAspGluValAlaGlnIleSerSerAspGluAlaIleGluThrAlaLys 259  
 DB 1165847 TTATCTATTATTATGCGCGTGAAGAACTGTTGATAGTATACCGCACTTGCCACAGCTGCT 1165906  
 QY 260 LeuLeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAla 279  
 DB 1165907 CGCTTAATGCGGAGAGAGAAATCTTGACAGGATTTCACTGTGTGACAGCTGTGCGGCT 1165966  
 QY 280 AlaPheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPhePro 299  
 DB 1165967 GCAGATCGCTTAGCTAAATTAACAGAAATTTGCTGATTAACCTCATTTGTTATTGCTT 1166026  
 QY 300 SerPheGlyGluArgTyrLeuSerSerValLeuPheGluSerVal 314  
 DB 1166027 TCAGCGTCTGAACGCTACTTAAGCACACGACTGTTTGAAGGATT 1166071

RESULT 11

US-09-643-990A-1  
 Sequence 1, Application US/09643990A  
 Patent No. 6528289  
 GENERAL INFORMATION:  
 APPLICANT: Robert D. Fleischmann  
 Mark D. Adams  
 Owen White  
 Hamilton O. Smith  
 J. Craig Venter  
 TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments thereof, and Uses thereof  
 NUMBER OF SEQUENCES: 1  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Human Genome Sciences, Inc.  
 STREET: 9410 Key West Avenue  
 CITY: Rockville,  
 STATE: MD  
 COUNTRY: USA  
 ZIP: 20850  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3 1/2 inch diskette  
 COMPUTER: Dell Pentium  
 OPERATING SYSTEM: MS DOS v6.22  
 SOFTWARE: ASCII Text  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/643,990A  
 FILING DATE: 23-Aug-2000  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/487,429  
 FILING DATE: 1995-06-07  
 APPLICATION NUMBER: 08/426,787  
 FILING DATE: 1995-04-21  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kenley K. Hoover  
 REGISTRATION NUMBER: 40,302  
 REFERENCE/DOCKET NUMBER: PB186P1C1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 301-610-5790  
 TELEFAX: 301-309-8439  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1830121 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-09-643-990A-1  
 Alignment Scores:  
 Pred. No.: 1.03e-81 Length: 1830121  
 Score: 816.50 Matches: 172  
 Percent Similarity: 70.16% Conservative: 49  
 Best Local Similarity: 54.60% Mismatches: 83



DB 4996 GCTAAGGATATAAGCATATTACTATGCCCCGAAACAATGAGCCAGACGCTGTAAT 5055  
QY 111 ILeuLeuAlaPheGlyAlaGluLeuValLeuThrAspProAlaLysGlyMetLysGly 130  
DB 5056 TTATTAAAGCATACCGTCCGGAATTAGTTTAACGCCCTGATCAGAAGCATGAAGGT 5115  
QY 131 AlaValGlnLysAlaGluGluLeuAlaLysThrProAlaLysTyrIle--LeuGln 149  
DB 5116 GCAATTAAAAAGCTTAAGAATTG-----AAGAGAGAACATGTTACTTCCAGCCACAA 5169  
QY 150 GlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThrGlyProGluIleTyr 169  
DB 5170 CAATTGAAACCCCTGCGAACCCTGAAGTTCACTTAAGTTAAGTTGTTGTTGTTGTTA 5229  
QY 170 LysGlySerAspGlyLys---IleAspAlaPheValSerGlyIleGlyThrGlyLysThr 188  
DB 5230 CAACAATTGAAGGGAATACTATCATGCGTCTAGCTGTTGTTGTTGTTGTTGTTGTTA 5289  
QY 189 IleThrGlyAlaGlyLysTyrLeuLysGluGlnAsnProAlaLysLysLysLysLysLys 208  
DB 5290 TTATCTGTTAGTAAAGTTCTGAAAAAAGAAATATCTTAACATCGAAATGTTGTTGTTA 5349  
QY 209 GluProValGluSerProValLeuSerGlyLysLysProGlyProHisLysIleGlnGly 228  
DB 5350 GAGCTGAGGCTTCTCCAGTATGAGCGGTGTTGAGCCAGTCCACATTAATTAACAAGGT 5409  
QY 229 IleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeuLeuAspGluValGln 248  
DB 5410 TTAGTGCTGATTTATTCAGGACCTTGAATACAGAAATCTATGACAGTATTATTAAA 5469  
QY 249 IleSerSerAspGluAlaIleGluThrAlaLysLeuLeuAlaLysGluGlyLeuPhe 268  
DB 5470 GTAGGAATGATACGCGATGGAATGCTCGTCAAGTTGCTAAAGAGAAAGTATTTTA 5529  
QY 269 ValGlyIleSerSerGlyAlaAlaAlaAlaAlaPheGlnIleAlaLysArgProGlu 288  
DB 5530 GCAGGTATTTCATCAGGTGCTCGATTATGCTCCATTCAAAAAGCAAGAA--TTA 5586  
QY 289 AsnAlaGlyLysLeuIleValAlaValPheProSerPheGlyLysArgLysSerSer 308  
DB 5587 GGAAGAGTAAACAGTAGTAACAGTATTGCGGAGTAATGCTGACGCTACTTATCAACA 5646  
QY 309 ValLeuPhe 311  
DB 5647 CCTTTATAT 5655  
RESULT 13  
US-09-107-532A-271  
Sequence 271, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107, 532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 271:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 927 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...927  
SEQUENCE DESCRIPTION: SEQ ID NO: 271:  
US-09-107-532A-271  
Alignment Scores:  
Pred. No.: 1.43e-86 Length: 927  
Score: 813.00 Matches: 171  
Percent Similarity: 70.83% Conservative: 50  
Best Local Similarity: 54.81% Mismatches: 85  
Query Match: 50.09% Indels: 6  
DB: 4 Gaps: 5  
US-09-931-457A-31 (1-325) x US-09-107-532A-271 (1-927)  
QY 5 ArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrProLeuValTyrLeu 24  
DB 1 AGAGACGCTTCTATTAAGTCACTACGACGAGTGTGTTGCAAAACGCCAATCGTCAATTA 60  
QY 25 AsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluPro 44  
DB 61 AATAAATCGTACACAGAGATTACAGATGTTTGTAAAGCTTGAAGTTTATATCA 120  
QY 45 CysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluLysGly 64  
DB 121 GAGAGAGTAAAGATCGTATGCTTTAAGTATGATAGAAAAGCAGAACACAGATGT 180  
QY 65 LeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIleGly 84  
DB 181 CTACTGAACACAGGTGATACATC---ATTGAACCTACTCTCGGAATAACAGGAATCGGA 237  
QY 85 LeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProAlaSerMet 104  
DB 238 TTGTGATGTAGAGAGTTCGCAAAAGATATAAGTAATCATCGTATGCTGAGACGATG 297  
QY 105 SerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspPro 124  
DB 298 AGTATCGAAGAAGCGTTTATTGATGAAGAAGGTACGCGCAGAGTGTGATTTGACTCTGT 357  
QY 125 AlaLysGlyMetLysGlyAlaValGlnLysAlaGluGluIleLeuAlaLysThrProAsn 144  
DB 358 GCGGATGTATTCTGATCATCAAGAGACAGACGT---TTGGCAAAAAGAAAACGCG 414  
QY 145 AlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThr 164  
DB 415 TACTTTTACCATTAACAATTGAAATAAGAAACAAATCTCTTGTCATGAGAAGACGACT 474  
QY 165 GlyProGluIleTyrLys-----GlySerAspGlyLysIleAspAlaPheValSerGly 182  
DB 475 GGTCTGAGATACATCAAGCATTTGAGTAATGCG---TTGACGCGTGTGTTGACAGT 531



183 ILeGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsnProAsn 202  
 532 ATCGGTACAGAGGAGCAATGATCAGTGGCGGCGGAGCAATTAAGCGTGTATCCGAAA 591  
 203 ILeLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLysProGly 222  
 592 ATCGAATTAATCGGGTAGAACACAGCAATCTGCTATTAGAGGAGGAGGAGGAGGAG 651  
 223 ProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeu 242  
 652 CCTCACAATTCAGAGCAATCGGTACAGGTTTCTTCTTCTTCTTCTTCTTCTTCTG 711  
 243 LeuAspGluValAlaGlnIleSerSerAspGluAlaIleGluThrAlaLysLeuAla 262  
 712 TATGATAAAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG 771  
 263 LeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAlaPheGln 282  
 772 CGAAAGAGAGAGATTTAGTACGATCTCATCTGGCGCTGCAATTGCTGCCGCTAAA 831  
 283 ILeAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSerPheGly 302  
 832 GTGCAAAAGAA---TTAGGAAAGGAGAGAGATTAAGTGTGTTCCGCAATGT 888  
 303 GluArgTyrLeuSerSerValLeuPheGluSerVal 314  
 889 GAACTTATTATTATCTACAGCCCTTACCAAGAGATA 924

RESULT 14

US-09-134-001C-493  
 ; Sequence 493, Application US/09134001C  
 ; Patent No. 6380370  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
 ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: GFC-007  
 ; CURRENT APPLICATION NUMBER: US/09/134,001C  
 ; CURRENT FILING DATE: 1998-08-13  
 ; PRIOR APPLICATION NUMBER: US 60/064,964  
 ; PRIOR FILING DATE: 1997-11-08  
 ; PRIOR APPLICATION NUMBER: US 60/055,779  
 ; PRIOR FILING DATE: 1997-08-14  
 ; NUMBER OF SEQ ID NOS: 5674  
 ; SEQ ID NO 493  
 ; LENGTH: 960  
 ; TYPE: DNA  
 ; ORGANISM: Staphylococcus epidermidis  
 ; US-09-134-001C-493

Alignment Scores:

Pred. NO.: 3.99e-85 Length: 960  
 Score: 801.00 Matches: 159  
 Percent Similarity: 72.52% Conservative: 60  
 Best Local Similarity: 52.65% Mismatches: 77  
 Query Match: 49.35% Indels: 6  
 DB: 4 Gaps: 5

US-09-931-457A-31 (1-325) x US-09-134-001C-493 (1-960)

QY 12 ValThrGluLeuIleGlyLysThrProLeuValTyrLeuAsnLysLeuAlaAspGlyCys 31  
 112 GCGAGCTGATATTATGTTAGTATGATAATCAAAATCCAGGTGTTGGTAAAGATCGT 171  
 DB 52 GTTACACAAATTAATGGAATACACTGTAATTAAGAAAGCTTGTGATGATGAT 111  
 QY 32 ValAlaArgValAlaAlaLysLeuGluLeuMetGluProCysSerSerValLysAspArg 51  
 112 GCGAGCTGATATTATGTTAGTATGATAATCAAAATCCAGGTGTTGGTAAAGATCGT 171  
 DB 52 GTTACACAAATTAATGGAATACACTGTAATTAAGAAAGCTTGTGATGATGAT 111  
 QY 52 IleGlyTyrSerMetIleAlaAspAlaGluGluLysGlyLeuIleThrProGlyLysSer 71  
 172 ATCGCTTTAGCGATGATGAAAAGCTGAGCGTGAAGGAGAAATTAACCTGCTGATACA 231  
 DB 72 ValLeuIleGluProThrSerGlyAsnThrGlyIleGlyLeuAlaPheMetAlaAlaAla 91

232 ATC---GTTAGCCTACAGAGTGATACACTGTATAGGCTTACGATTTGTATGTGCTGCC 288  
 QY 92 ArgGlyTyrLysLeuIleIleThrMetProAlaSerMetSerLeuGluArgGlyIle 111  
 289 AAGGGTACAAAGCAGTTTACATGCTGAAACAAATGAGCCAGAGCGCTTACTTA 348  
 QY 112 LeuLeuAlaPheGlyAlaGluLeuValLeuThrAspProAlaLysGlyMetLysGlyAla 131  
 349 TTAAAGCTTATGCTGCTGAACTAGTATTAACACAGGATCTGAAGCTATGAAGGTGCA 408  
 QY 132 ValGlnLysAlaGluGluIleLeuAlaLysThrProAsnAlaTyrIle---LeuGlnGln 150  
 409 ATAAAAAGCTAAAGATTA-----AAGAGAGACACGGCTATTGTAACCAACAA 462  
 QY 151 PheGluAsnProAlaAsnProLysValHisTyrGluThrThrGlyProGluIleTyrLys 170  
 463 TTGAAATTCAGCAAAATCTGAAATCATGAACCTTACAACTGGAACCAAGATTAGTTGAA 522  
 QY 171 GlySerAspGlyLys---IleAspAlaPheValSerGlyIleGlyThrGlyThrIle 189  
 523 CAATTGAGGTGACAAATGATGATTTTAACTGTGTAGAGAACTGGTGTACGTTA 582  
 DB 190 ThrGlyAlaGlyLysTyrLeuLysGluGlnAsnProAsnIleLysLeuIleGlyValGlu 209  
 583 TCTGTGTGTGTAAGTATTAAGAGAAAGATATCCAAATGTGAAATAGTACTATTGAA 642  
 QY 210 ProValGluSerProValLeuSerGlyLysProGlyProHisLysIleGlnGlyIle 229  
 643 CCTGAAGCTTCTCCAGTATTAAAGCGGTGTGAACCAAGCCCTCATTAATTAAGAGATTG 702  
 DB 230 GlyAlaGlyPheIleProGlyValLeuGluValAsnLeuLysAspGluValAlaGlnIle 249  
 703 GAGCAGGTTTCTGTAACCTGATCTTAAATACAGAAAGTTATGACAGCATCAAAAGTA 762  
 QY 250 SerSerAspGluAlaIleGluThrAlaLysLeuLeuAlaLysGluGlyLeuPheVal 269  
 763 GGTATGATAGTCTGATGATATATGACAGCTGTGTTCTAGAGAGAGGATATTATAGCA 822  
 DB 270 GlyLysSerSerGlyAlaAlaAlaAlaAlaAlaPheGlnIleAlaLysArgProGluAsn 289  
 823 GGTATTTCATCTGCTGCTGCAATATATATGCTCTATTCAAAAAGCAAAAGAA---TTAGGT 879  
 QY 290 AlaGlyLysLeuIleValAlaValAlaPheProSerPheGlyGluArgTyrLeuSerVal 309  
 880 AAAGTAAACAGTTGTAAACAGTATTACCAAGTATGAGGAAAGCTTACTTATCAACACCA 939  
 DB 310 LeuPhe 311  
 940 TTATAT 945

RESULT 15

US-08-961-527-15  
 ; Sequence 15, Application US/08961527  
 ; Patent No. 6420135  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Charles Kunsch  
 ; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESS: Human Genome Sciences, Inc.  
 ; STREET: 9410 Key West Avenue  
 ; CITY: Rockville  
 ; STATE: Maryland  
 ; COUNTRY: USA  
 ; ZIP: 20850  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage  
 ; COMPUTER: HP Vectra 486/33  
 ; OPERATING SYSTEM: MSDOS version 6.2  
 ; SOFTWARE: ASCII Text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/961,527



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 21, 2004, 23:43:26 ; Search time 577 seconds  
(without alignments)  
2559.855 Million cell updates/sec

Title: US-09-931-457A-31  
Perfect score: 1623  
Sequence: 1 MAVERSGIADVTBLIKTP.....LSSVPRSVRRAESMTFEP 325

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2953838 segs, 2272363821 residues

Total number of hits satisfying chosen parameters: 5907676

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09931457/runat\_18052004\_121731\_20391/app\_query.fasta\_1.519  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=Dct -THR MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09931457 @CGN 1 1 57 @runat\_18052004\_121731\_20391  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

Result No. Score Query Match length DB ID

Description

## SUMMARIES

1	1623	100.0	1362	9	US-09-931-457A-30	Sequence 30, Appl
2	1622	99.9	1407	13	US-10-424-599-130460	Sequence 130460, A
3	1488	91.7	3075	13	US-10-424-599-55164	Sequence 55164, A
4	1389	85.6	1319	13	US-10-425-114-25435	Sequence 25435, A
5	1381	85.1	1308	13	US-10-425-114-19467	Sequence 19467, A
6	1381	85.1	1403	13	US-10-425-114-1445	Sequence 1445, App
7	1380	85.0	1312	13	US-10-425-114-23298	Sequence 23298, A
8	1375	84.7	1301	13	US-10-425-114-23608	Sequence 23608, A
9	1374	84.7	966	9	US-09-938-842A-693	Sequence 893, App
10	1374	84.7	966	11	US-09-938-842A-893	Sequence 893, Appl
11	1317	81.1	1522	9	US-09-931-457A-63	Sequence 63, Appl
12	1316	81.1	1404	13	US-10-424-599-55160	Sequence 55160, A
13	1309	80.7	1279	13	US-10-425-114-22525	Sequence 22525, A
14	1279	78.8	1490	13	US-10-425-114-34372	Sequence 34372, A
15	1279	78.8	1491	13	US-10-425-114-34527	Sequence 34527, A
16	1275	78.6	1508	9	US-09-931-457A-61	Sequence 61, Appl
17	1275	78.6	1509	13	US-10-425-114-30493	Sequence 30493, A
18	1255	77.3	1329	9	US-09-938-842A-99	Sequence 99, Appl
19	1255	77.3	1329	11	US-09-938-842A-99	Sequence 99, Appl
20	1234	76.0	1603	13	US-10-424-599-61234	Sequence 61234, A
21	1224	75.4	2307	13	US-10-425-114-7548	Sequence 7548, App
22	1192	73.4	972	9	US-09-938-842A-495	Sequence 495, App
23	1192	73.4	972	11	US-09-938-842A-495	Sequence 495, App
24	1192	73.4	1472	9	US-09-887-576-611	Sequence 611, App
25	1179	72.6	975	10	US-09-905-290A-3	Sequence 3, Appl
26	1079	66.5	1563	13	US-10-424-599-53377	Sequence 53377, A
27	968.5	59.7	945	13	US-10-282-122A-12236	Sequence 12236, A
28	949.5	58.5	930	13	US-10-282-122A-26520	Sequence 26520, A
29	949.5	58.5	933	13	US-10-282-122A-28508	Sequence 28508, A
30	947.5	58.4	1925	13	US-10-424-599-18037	Sequence 18037, A
31	941.5	58.0	930	13	US-10-282-122A-25894	Sequence 25894, A
32	941.5	58.0	1086	13	US-10-282-122A-26080	Sequence 26080, A
33	940.5	57.9	933	13	US-10-282-122A-27578	Sequence 27578, A
34	917	56.5	927	13	US-10-282-122A-26908	Sequence 26908, A
35	917	56.5	100848	13	US-10-672-787-39	Sequence 39, Appl
36	916.5	56.5	927	13	US-10-282-122A-16019	Sequence 16019, A
37	912	56.2	933	13	US-10-282-122A-29654	Sequence 29654, A
38	897	55.3	588	16	US-10-341-961A-271	Sequence 271, App
39	896.5	55.2	930	13	US-10-282-122A-15848	Sequence 15848, A
40	882	54.3	969	13	US-10-282-122A-41010	Sequence 41010, A
41	879	54.2	951	13	US-10-282-122A-32551	Sequence 32551, A
42	870	53.6	996	13	US-10-282-122A-8820	Sequence 8820, App
43	865.5	53.3	972	13	US-10-282-122A-38798	Sequence 38798, A
44	865.5	53.3	972	13	US-10-282-122A-39402	Sequence 39402, A
45	863.5	53.2	969	13	US-10-282-122A-19471	Sequence 19471, A

## ALIGNMENTS

RESULT 1  
US-09-931-457A-30  
Sequence 30, Application US/09931457A  
Patent No. US20020157132A1  
GENERAL INFORMATION:  
APPLICANT: Falco, S. Carl  
TITLE OF INVENTION: Plant Amino Acid Biosynthetic Enzymes  
FILE REFERENCE: B1116 US CIP  
CURRENT APPLICATION NUMBER: US/09/931,457A  
CURRENT FILING DATE: 2002-02-22  
PRIOR APPLICATION NUMBER: 09/424,976  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: 60/065,385  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/049,406  
PRIOR FILING DATE: 1997-06-12  
NUMBER OF SEQ ID NOS: 72  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 30  
LENGTH: 1362  
TYPE: DNA  
ORGANISM: Glycine max

US-09-931-457A-30

## Alignment Scores:

Pred. No.:	1,84e-178	Length:	1362
Score:	1623.00	Matches:	325
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-931-457A-31 (1-325) x US-09-931-457A-30 (1-1362)

```

QY 1 MetAlaValGluArgSerGlyIleAlaIysAspValThrGluLeuIleGlyLysThrPro 20
DB 90 ATGGCTGTGAAGGTCCTGGAATGCGAAGATGCTTACGGAATGATGGTAAACCCCA 149
QY 21 LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40
DB 150 TTAGTATATCTAAATAAACTTGGGATGTTGTGTGCGCGGTGCTGCTAAACTGAG 209
QY 41 LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60
DB 210 TTGATGAGCCATGCTCTAGTGTGAAGCAGAGATTGGTATAGTATGATGCTGATGCA 269
QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80
DB 270 GAAGAGAGAGGACTTATCACAACCTGGAAGAGTGCTCTCATTTGAGCCACAAGTGTAT 329
QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMet 100
DB 330 ACTGGCATTTGATTAAGCTTTCATGCGCAGCAGCCAGGGTTACAGCTCATATTACATG 389
QY 101 ProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120
DB 390 CCTGCTTTATGATGCTTTGAGAGAAAGATCATTTCTATTAGCTTTTGAGAGCTGAGTTG 449
QY 121 LeuThrAspProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluGluIleLeuAla 140
DB 450 CTGACAGATCTGCTTAAGGAAATGAAGGTGCTGTTCAAGAGGCTGAAGAGATATTGCT 509
QY 141 LysThrProAsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHis 160
DB 510 AAGAGCCCAATGCTCTACATCTTCAACATTTGAAAAACCTGCTCAATCCCAAGGTTCT 569
QY 161 TyrGluThrThrGlyProGluIleTrrLysGlySerAspGlyLysIleAspAlaPheVal 180
DB 570 TATGAAACCACTGCTGCTCAGAGATATGAAAGGCTCGATGGGAAATTTGATGCTTTG 629
QY 181 SerGlyIleGlyThrGlyGlyTyrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsn 200
DB 630 TCTGGATAGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 689
QY 201 ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyGlyLys 220
DB 690 CCGAATATTAAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 749
QY 221 ProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240
DB 750 CCTGCTCACAAGAGATTCAAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 809
QY 241 AsnLeuLeuAspGluValAlaGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260
DB 810 AATCTTCTTGAATGAGTTGTTCAATATCAAGTGAAGCAATGAAATGCAAAAGCTT 869
QY 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyValAlaAlaAlaAla 280
DB 870 CTTCGCTTAAAGAGGCTTATTTGTGGGATATCTTCCGAGCTGAGCTGCTGCTGCT 929
QY 281 PheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSer 300
DB 930 ETTCAAGATTGAAAAAGACCAAGAAATGCGGGAAGCTTATTTGCGGTTTTCACAGC 989
QY 301 PheGlyLysArgTyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSer 320

```

```

DB 990 TTCGGGAGAGGTACTCTGCTCCGTGCTATTGAGTCACTGAGACGCGAAGCTGAAAGC 1049
QY 321 MetThrPheGluPro 325
DB 1050 ATGACTTTTGAGCCC 1064

```

## RESULT 2

US-10-424-599-130460

```

; Sequence 130460, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 130460
; LENGTH: 1407
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_88813C.1
US-10-424-599-130460

```

## Alignment Scores:

Pred. No.:	2.52e-178	Length:	1407
Score:	1622.00	Matches:	324
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.69%	Mismatches:	0
Query Match:	99.94%	Indels:	0
DB:	13	Gaps:	0

US-09-931-457A-31 (1-325) x US-10-424-599-130460 (1-1407)

```

QY 1 MetAlaValGluArgSerGlyIleAlaIysAspValThrGluLeuIleGlyLysThrPro 20
DB 126 ATGGCTGTGAAGGTCCTGGAATGCGAAGATGCTTACGGAATGATGGTAAACCCCA 185
QY 21 LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40
DB 186 TTAGTATATCTAAATAAACTTCCGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 245
QY 41 LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60
DB 246 TTGATGAGCCATGCTCTAGTGTGAAGCAGAGATTGGTATAGTATGATGCTGATGCA 305
QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80
DB 306 GAAGAGAGGACTTATCACAACCTGGAAGAGTGCTCTCATTTGAGCCACAAGTGTAT 365
QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMet 100
DB 366 ACTGCAATTTGATTAAGCTTCAATGCGCAGCAGCCAGGGTTACAGCTCATTAATTACAT 425
QY 101 ProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120
DB 426 CTTGCTTTCTATGAGTCTTGAAGAGAAATCATTTCTATTAGCTTTTGAGGCTGAGTTG 485
QY 121 LeuThrAspProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluGluIleLeuAla 140
DB 486 CTGACAGATCTGCTTAAGGAAATGAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 545
QY 141 LysThrProAsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHis 160
DB 546 AAGAGCCCAATGCTCTACATCTTCAACATTTGAAAAACCTGCTCAATCCCAAGGTTCT 605
QY 161 TyrGluThrThrGlyProGluIleTrrLysGlySerAspGlyLysIleAspAlaPheVal 180

```



|||||  
Db 606 TATGAAACCACTGGTCCAGAGATATGAAAGGCTCCGATGGGAAATATGATGATTGTT 665  
QY 181 SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsn 200  
|||  
Db 666 TCTGGATAGGCACTGCTGTGTAACAATAACAGGTGCTGAAATAATCTTAAGAGCAGAAAT 725  
QY 201 ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLys 220  
|||  
Db 726 CCGATATATAAGCTGATGGTGGAAACAGTTGAAAGTCCAGTCTCAGAGGAAAG 785  
QY 221 ProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240  
|||  
Db 786 CCTGGTCCACACAGATTCAGAGGATGGTGTCTGTTATCCCTGTGTCTTGGAAATC 845  
QY 241 AsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260  
|||  
Db 846 AATCTTCTTGATGAGTATATCAATATCAAGTGAAGCAATAGAACTGCAAGCTT 905  
QY 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla 280  
|||  
Db 906 CTTCGGCTTAAGAAGGCTATATTGTGGAAATATCTCCGAGAGCTGACGTGCTGCT 965  
QY 281 PheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSer 300  
|||  
Db 966 TTTTCAGATTGCAAAAAGACCAAGAAATCCCGGAGCTTATTGTTCCTTTTCCAGC 1025  
QY 301 PheGlyGluArgTyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSer 320  
|||  
Db 1026 TTGGGGAGAGGTACTGCTCTCTCTGCTATTGAGTCAAGAGAGCGGAGCTGAAAGC 1085  
QY 321 MetThrPheGluPro 325  
|||  
Db 1086 ATGACTTTTGAGCCC 1100

RESULT 3  
US-10-424-599-55164  
; Sequence 55164, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 55164  
; LENGTH: 3075  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURES:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_20824C.1  
US-10-424-599-55164

Alignment Scores: 2.94e-162 Length: 3075  
Pred. No.: 1488.00 Matches: 290  
Score: 1488.00 Conservative: 21  
Percent Similarity: 95.69% Mismatches: 14  
Best Local Similarity: 89.23% Indels: 0  
Query Match: 91.68% Gaps: 0  
DB: 13

US-09-931-457A-31 (1-325) X US-10-424-599-55164 (1-3075)

QY 1 MetAlaValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrPro 20  
|||  
Db 312 ATGGCGGTGAGAGTTGACATTGCAAAAAGATGAGCTGAATTAATTGCAAAAACCTCCA 371  
QY 21 LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40

|||||  
Db 372 TTGGTATATCTCAACCGCATCGTGATGGCTGTGTGTCACAAAGTTGTCGGAAGCTAGAA 431  
QY 41 LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60  
|||  
Db 432 ATGATGAACCTTGCTCCAGTGTCAAAAGACAGATAGGGTATAGCATGATTGTAGATGCT 491  
QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80  
|||  
Db 492 GAGAGAAAGGACTCATCACACCGGTGAAGTGTCTCATTTAGAGCTTACTAGTGAAG 551  
QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrTyrLysLeuIleThrMet 100  
|||  
Db 552 ACTGCATAGGTTGGCATTCATGACAGCTGCTAAGGTTATATAACTTATATAACCATG 611  
QY 101 ProAlaSerMetSerLeuGluArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120  
|||  
Db 612 CCTTCTTGATGAGTCTCGAAAGAGAACCATCTTCGAGCTTTGAGCTGAGTTAGTC 671  
QY 121 LeuThrAspProAlaLysGlyMetLysGlyAlaValGlnLysAlaGlnIleLeuAla 140  
|||  
Db 672 CTCAGTATCCCGCAAGATATGAAGAGCTGTTCAAGAGCAGAGAGATTCGGAT 731  
QY 141 LysThrProAsnAlaTyrIleLeuGlnIlePheGluAsnProAlaAsnProLysValHis 160  
|||  
Db 732 AAGACTCCAAATTTCTATATCTTCAGCAATTTGAAAATCTGCAACCCCAAGATCCAT 791  
QY 161 TyrGluThrThrGlyProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheVal 180  
|||  
Db 792 TATGAACCACTGACCGAGATCTGGAAGGCTCTAGTGGAGAGGTGATGCTCTTGTG 851  
QY 181 SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsn 200  
|||  
Db 852 TCTGGATTGGACCTGAGAGTACATTAACAGTCTGGAATAATATCTGAAGATCAGAAC 911  
QY 201 ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLys 220  
|||  
Db 912 CCTGATATTAAGCTATATGTCATTCAGCAAGTGAAGAGTCCAAATCTGCTGAGGAAA 971  
QY 221 ProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240  
|||  
Db 972 CCTGGCCCTCAATAGATCCAGGAATGGTGTGCTTCATCCCTGCTGTGATGTC 1031  
QY 241 AsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260  
|||  
Db 1032 GATTACTTGATGAAGTTGTTCAATTCAACTGAAGAGCCATTGAACCTGAACCT 1091  
QY 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla 280  
|||  
Db 1092 CTTCCTTGAAGAAGGTTGTGCTGGGAATATCATCTGTGCACTGCACTGCTGCA 1151  
QY 281 PheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSer 300  
|||  
Db 1152 ATTAAGATGCAAAAGAGCCAGAAATGCTGAAAACCTCATTTGCGCTTCCCAAGC 1211  
QY 301 PheGlyGluArgTyrLeuSerSerValLeuPheGluSerValArgGluAlaGluSer 320  
|||  
Db 1212 TTTGAGAGGCTTATTTATCATCTGCTCTTTGAATCCGTGAAGCAGAGAGAGAGC 1271  
QY 321 MetThrPheGluPro 325  
|||  
Db 1272 ATGCTCTTCAAGCCT 1286

RESULT 4  
US-10-425-114-25435  
; Sequence 25435, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jindong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack B

APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(53313) B  
 CURRENT APPLICATION NUMBER: US/10/425,114  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 73128  
 SEQ ID NO 25435  
 LENGTH: 1319  
 TYPE: DNA  
 ORGANISM: Zea mays  
 FEATURE:  
 OTHER INFORMATION: Clone ID: LIB3900-005-C7\_FLI  
 US-10-425-114-25435

Alignment Scores:  
 Pred. No.: 2,76e-151 Length: 1319  
 Score: 1389.00 Matches: 270  
 Percent Similarity: 91.82% Conservative: 22  
 Best Local Similarity: 84.91% Mismatches: 26  
 Query Match: 85.58% Indels: 0  
 DB: 13 Gaps: 0

US-09-931-457A-31 (1-325) x US-10-425-114-25435 (1-1319)

QY 8 IleAlaLysAspValThrGluLeuIleGlyLysThrProLeuValTyrLeuAsnLysLeu 27  
 DB 112 ATGCCCAAGAGCGTCAACCGAGTGTATCGGGAACACCGCGCTGTGTACTTCAACAGGTG 171  
 QY 28 AlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluProCysSerSer 47  
 DB 172 ACCGATGGTGGTGGCGCGCGCTGCGCGCAAGCTCGAGTCCATGAGCCCTGCTCCAGC 231  
 QY 48 ValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluLysGlyLeuIleThr 67  
 DB 232 GTCAAGATAGATGGCTTACATGATGACGACGACGACGACGACGACGACGACGACGACGAC 291  
 QY 68 ProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIleGlyLeuAlaPhe 87  
 DB 292 CCAGCGGTGAGTGTCTGATTTGAACCACTACCGGCAACAGGCAATGAGCTGCTTT 351  
 QY 88 MetAlaAlaAlaArgGlyTyrLysLeuIleThrMetProAlaSerMetSerLeuGlu 107  
 DB 352 ATGGCTGCTGCCAAGGCTCAACAATTAACCTACATCAACAATGCTGCTCAATGAGCATGAG 411  
 QY 108 ArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspProAlaLysGly 127  
 DB 412 AGGAGATCATATGTAAGGCTTTGTGTGCTGTAAGCTGTCTTACTGACCACTCTTGGA 471  
 QY 128 MetLysGlyAlaValGlnLysAlaGluGluIleLeuAlaLysThrProAsnAlaTyrIle 147  
 DB 472 ATGAAGAGAGCTGTCAAGAAAGCGAAGAGATACAGCAAGAACACCACTCGTACATT 531  
 QY 148 LeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThrGlyProGlu 167  
 DB 532 CTTCACAAATTTGAATCCAGCTAACCCCAAGATTCATATGAGACTACGCGCTGAA 591  
 QY 168 IleTyrLysGlySerAspGlyLysIleAspAlaPheValSerGlyIleGlyThrGlyGly 187  
 DB 592 ATCTGGAAGCTACAGCAAGAAATGATGCTGTATCTGTATCTGATCGGACAGAGCT 651  
 QY 188 ThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsnProAsnIleLysLeuIleGly 207  
 DB 652 ACCATCACCGGTACTGCGGCGATACCTCAGAGCAAAATCTATATGTCAAGCTCTATGCT 711  
 QY 208 ValGluProValGluSerProValLeuSerGlyGlyLysProGlyProHisLysIleGln 227  
 DB 712 GTGAGCCAGGTGAGAGTCTGTTTGAATGTGAGAAACCTGAGCAACCAAGATTCAA 771  
 QY 228 GlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeuLeuAspGluValVal 247  
 DB 772 GGAATTGAGAGCTGTTTATCCCTGAGTCTTGAATGTGATCTCTTGAATGAACTCTA 831

QY 248 GlnIleSerSerAspGluAlaIleGluThrAlaLysLeuLeuAlaLeuLysGluGlyLeu 267  
 DB 832 CAGTTTCAAGTATGATGAGCTATCGAAGCTGCGCAAGGCTCTTCTGAAAGAGGTTG 891  
 QY 268 PheValGlyIleSerSerGlyAlaAlaAlaAlaAlaPheGlnIleAlaLysArgPro 287  
 DB 892 TTGTTGAAATCTTCTGTCAGCTGACGTCGAGCTGCGGAGTTAGCTTGAAGAGCCA 951  
 QY 288 GluAsnAlaGlyLysLeuIleValAlaValPheProSerPheGlyGluArgTyrLeuSer 307  
 DB 952 GAAACGCGCGGAAAGCTATTGTTGTGCTGTTCCGAGCTTCCGAGCGCTACTCTCA 1011  
 QY 308 SerValLeuPheGluSerValArgGlyAlaGluSerMetThrPheGluPro 325  
 DB 1012 TCGGTGCTGTTCAGTCCATCAAGAGAGCGGAGAGCATGTGTGCGAGCCC 1065

RESULT 5

US-10-425-114-19467  
 Sequence 19467, Application US/10425114  
 Publication No. US20040034888A1  
 GENERAL INFORMATION:  
 APPLICANT: Lin, Jingdong  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Screen, Steven B.  
 APPLICANT: Tabaska, Jack B.  
 APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53313) B  
 CURRENT APPLICATION NUMBER: US/10/425,114  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 73128  
 SEQ ID NO 19467

LENGTH: 1308

TYPE: DNA

ORGANISM: Zea mays

FEATURE:  
 OTHER INFORMATION: Clone ID: LIB3150-027-DB\_FLI

US-10-425-114-19467

Alignment Scores:  
 Pred. No.: 2.32e-150 Length: 1308  
 Score: 1381.00 Matches: 269  
 Percent Similarity: 91.51% Conservative: 22  
 Best Local Similarity: 84.59% Mismatches: 27  
 Query Match: 85.09% Indels: 0  
 DB: 13 Gaps: 0

US-09-931-457A-31 (1-325) x US-10-425-114-19467 (1-1308)

QY 8 IleAlaLysAspValThrGluLeuIleGlyLysThrProLeuValTyrLeuAsnLysLeu 27  
 DB 101 ATGCCCAAGAGCGTCAACCGAGTGTATCGGGAACACCGCGCTGTGTACTTCAACAGGTG 160  
 QY 28 AlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluProCysSerSer 47  
 DB 161 ACCGATGGTGGTGGCGCGCGCTGCGCGCAAGCTCGAGTCCATGAGAGCTGCTCCAGC 220  
 QY 48 ValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluLysGlyLeuIleThr 67  
 DB 221 GTCAAGATAGATTTGCTGATGAAACCACTAGCGGCAACAGGATTTGACTGTA 280  
 QY 68 ProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIleGlyLeuAlaPhe 87  
 DB 281 CCAAGCGTGAAGTGTCTGATGAAACCACTAGCGGCAACAGGATTTGACTGTA 340  
 QY 88 MetAlaAlaAlaArgGlyTyrLysLeuIleThrMetProAlaSerMetSerLeuGlu 107  
 DB 341 ATGGCTCTGCCAAGGCTCAACAATTAACCTACATCAACAATGCTGCTCCATGAGCATGAG 400  
 QY 108 ArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspProAlaLysGly 127

DB 401 AGGAGATCATATGAAAGCTTTTGTGCTGAACTGTCTTACTGACCCACTCCTGGGA 460  
 QY 128 MetLysGlyAlaValGlnLysAlaGluGluLeuAlaLysThrProAsnAlaTyrIle 147  
 DB 461 ATGAAAGAGCTGTCAAGAAAGCGAGAGATACAGCAAGACACCCACTCTACATC 520  
 QY 148 LeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThrGlyProGlu 167  
 DB 521 CTCACAAATTGAAAATCCAGCTAACCCAAAGATCACTATGAGACTACTGGCCTGAA 580  
 QY 168 IleTyrLysGlySerAspGlyLysIleAspAlaPheValSerGlyIleGlyThrGlyGly 187  
 DB 581 ATCTGAAAGCTACAGCAGAAATAATGCTGGCCTTGTATCTGATATCGGAGACAGAGGT 640  
 QY 188 ThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsnProAsnIleLysLeuIleGly 207  
 DB 641 ACCATCACCCTGCTACTGGCGGATACCTCAGAGCAAAATCTAATGCTCAAGCTCTATGCT 700  
 QY 208 ValGluProValGluSerProValLeuSerGlyGlyLysProGlyProHisLysIleGln 227  
 DB 701 GTGAGCCAGTGAAGTGTCTGTTTGAATGTGGAATACTGAGCCACACAGATTCAA 760  
 QY 228 GlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeuLeuAspGluValVal 247  
 DB 761 GGAATTGAGCTGTTTATCCCTGAGTCTTGATGATCTCTGATGAGACTCTA 820  
 QY 248 GlnIleSerSerAspGluAlaIleGluThrAlaLysLeuLeuAlaLeuLysGluGlyLeu 267  
 DB 821 CAGTTTCAAGTGAAGCTATGTAGACTGCAAGCTCTTGTCTGAAAGAGGTTG 880  
 QY 268 PheValGlyIleSerSerGlyAlaAlaAlaAlaAlaPheGlnIleAlaLysArgPro 287  
 DB 881 TTGTTGGAATCTTCTGCTGAGCTGAGCTGCGGAGTGAAGCTTGAAGAGCCA 940  
 QY 288 GluAsnAlaGlyLysLeuIleValAlaValPheProSerPheGlyGluArgTyrLeuSer 307  
 DB 941 GAAACGCCGGAAGCTATTGTTGCTGCTGCTCCGAGCTTCCGCGAGGCTACCTCTCA 1000  
 QY 308 SerValLeuPheGluSerValArgArgGluAlaGluSerMetThrPheGluPro 325  
 DB 1001 TCGGTCTGTTCCAGTCCATCAAGAGGAGCGGAAAGCATGTGTGTCAGCCC 1054

RESULT 6

US-10-425-114-1445  
 ; Sequence 1445, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jindong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(5313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 1445  
 ; LENGTH: 1403  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: 700153393\_FLI  
 ; US-10-425-114-1445

Alignment Scores:  
 Pred. No.: 2.57e-150 Length: 1403  
 Score: 1381.00 Matches: 269  
 Percent Similarity: 91.51% Conservative: 22  
 Best Local Similarity: 84.59% Mismatches: 27  
 Query Match: 85.09% Indels: 0

DB: 13 Gaps: 0

US-09-931-457a-31 (1-325) x US-10-425-114-1445 (1-1403)

QY 8 IleAlaLysAspValThrGluLeuIleGlyLysThrProLeuValTyrLeuAsnLysLeu 27  
 DB 196 ATGCCAAGAGAGCTACCGAGTGTATCGGGAACACCGCGCTGTGTACCTCAACAAGGTG 255  
 QY 28 AlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluProCysSerSer 47  
 DB 256 ACCGATGGGTGCTCGGCGCGCTCGCCCAAGCTCGAGTCCATGGAAGCCCTGCTCAGC 315  
 QY 48 ValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGluLysGlyLeuIleThr 67  
 DB 316 GTCAAGATGAGATTGGCTACAGCATGATCAAGGACGAGAGAGAAAGGCTGATTACT 375  
 QY 68 ProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIleGlyLeuAlaPhe 87  
 DB 376 CCAGGCGTGAAGTCTTCTGATTGAACCACTAGCGGCAACAGGCAATTGAGCTGGCTTT 435  
 QY 88 MetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProAlaSerMetSerLeuGlu 107  
 DB 436 ATGGCTGCTCCAAAGGCTACAACTTACACTCACAATGCTGCTCCTCATGAGCATGAG 495  
 QY 108 ArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspProAlaLysGly 127  
 DB 496 AGGAGATCATATGAAAGCTTTTGTGCTGAACTTGCTTACTGACCCACTCCTGGGA 555  
 QY 128 MetLysGlyAlaValGlnLysAlaGluGluIleLeuAlaLysThrProAsnAlaTyrIle 147  
 DB 556 ATGAAAGAGCTGTCAAGAAAGCGAAGAGATACAAGCAAGACACCACTCGTACATC 615  
 QY 148 LeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThrGlyProGlu 167  
 DB 616 CTTCACAAATTGAAAATCCAGCTTAACCAAGATTCACTAAGACTACTGGGCTGAA 675  
 QY 168 IleTyrLysGlySerAspGlyLysIleAspAlaPheValSerGlyIleGlyThrGlyGly 187  
 DB 676 ATCTGAAAGCTACAGCAGAAATAATGCTGCGCTTGTATCTGTATCGGAGCAGAGGT 735  
 QY 188 ThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsnProAsnIleLysLeuIleGly 207  
 DB 736 ACCATCACCCTGCTACTGGCGGATACCTCAGAGCAAAATCTAATGCTCAAGCTCTATGCT 795  
 QY 208 ValGluProValGluSerProValLeuSerGlyGlyLysProGlyProHisLysIleGln 227  
 DB 796 GTGAGCCAGTGAAGAGTGTCTTGAATGTGGAATACTGAGCAACCAAGATTCAA 855  
 QY 228 GlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeuLeuAspGluValVal 247  
 DB 856 GGAATTGAGCTGTTTATCTGCTGAGTCTTGATGTGATCTCTGATGAGACTCTTA 915  
 QY 248 GlnIleSerSerAspGluAlaIleGluThrAlaLysLeuLeuAlaLeuLysGluGlyLeu 267  
 DB 916 CAGTTTCAAGTGAAGCTATGTAGACTGCCAAGGCTCTGCTCTGAAAGAGGTTG 975  
 QY 268 PheValGlyIleSerSerGlyAlaAlaAlaAlaAlaPheGlnIleAlaLysArgPro 287  
 DB 976 TTGTTGGAATCTTCTGCTGAGCTGCAAGCTGCGGAGTGAAGCTTGTCTTAAGAGCCA 1035  
 QY 288 GluAsnAlaGlyLysLeuIleValAlaValPheProSerPheGlyGluArgTyrLeuSer 307  
 DB 1036 GAAACGCCGGAAGCTATTGTTGCTGCTGCTCCGAGCTTCCGCGAGGCTACCTCTCA 1095  
 QY 308 SerValLeuPheGluSerValArgArgGluAlaGluSerMetThrPheGluPro 325  
 DB 1096 TCGGTCTGTTCCAGTCCATCAAGAGGAGCGGAAAGCATGTGTGTCAGAGCCC 1149

RESULT 7

US-10-425-114-23298  
 ; Sequence 23298, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:



APPLICANT: Liu, Jingdong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E.  
APPLICANT: Tabaska, Jack E.  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 23298  
LENGTH: 1312  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB3597-017-F4\_FLI  
US-10-425-114-23298

Alignment Scores:  
Pred. No.: 3.05e-150 Length: 1312  
Score: 1380.00 Matches: 269  
Percent Similarity: 91.51% Conservative: 22  
Best Local Similarity: 84.59% Mismatches: 27  
Query Match: 85.03% Indels: 0  
DB: 13 Gaps: 0

US-09-931-457a-31 (1-325) x US-10-425-114-23298 (1-1312)

QY 8 TLEAlAlAspValThrgLUleuilegLYsThrProleuValTyrlEuAsnLYsLeu 27  
DB 105 ATGCCAAGAGAGCTGACCGAGTGTATCGGACACGCCCTGTGTACTTCAACAAGTG 164  
QY 28 AlaAspGlyCyValAlaArgValAlaAlaLYsLeuGluMetGluProCySerSer 47  
DB 165 ACCGATGGTGGTGGCGCGCGCTGCCCGCAAGCTCGAGTCCATGAGGCCCTGCTCAGC 224  
QY 48 ValLYsAspArgilegLYsSerMetileAlaAspAlaGluGluLYsGlyLeuileThr 67  
DB 225 GTCAAGATAGATTGGCTACAGCATGATCAGGACGACAGAGAGAGAGGCTGATTACT 284  
QY 68 ProGlyLYsSerValleuilegLUProThrSerGlyAsnThrgLYleuAlaPhe 87  
DB 285 CCAAGCGTGAAGTGTCTGATTGAACCAACTAGCGCAACAGGCATTGGACTGGCTTT 344  
QY 88 MetAlaAlaArgGlyTyrlLYsLeuileleThrMetProAlaSerMetSerLeuGlu 107  
DB 345 ATGGCTGCTGCCAAGGCTACAACTTACACTCAATGCTGCTCCATGACATGAG 404  
QY 108 ArgArgileleleuAlaPheGlyAlaGluLeuValleuThrAspProAlaLYsGly 127  
DB 405 AGAGGATCATATTGAAGCTTTGGTGTGAACCTTACTTACTGACCACTCTTGGGA 464  
QY 128 MetLYsGlyAlaValGlnLYsAlaGluGluileuAlaLYsThrProAlaAlaTyrlle 147  
DB 465 ATGAAGAGAGCTGTACAGAAAGCGAGAGATACAGCAAGACACCACTCGTACATT 524  
QY 148 LeuGlnGlnPheGluAsnProAlaAsnProLYsValHisTYrGluThrThrgLYProGlu 167  
DB 525 CTTCACAAATTGAATAATCCACTAACCCAAAGATTCACTGTAGACTACTGGGCTGAA 584  
QY 168 TLeTrpLYsGlySerAspGlyLYsileAspAlaPheValSerGlyileGlyThrgLYGly 187  
DB 585 ATCTGAAGAGCTACAGCAGAAATAATGATGCTGTATCTGTATCGGACAGAGAGT 644  
QY 188 ThrileThrgLYAlaGlyLYsTYrleuLYsGluGlnAsnProAsnileLYsleuileGly 207  
DB 645 ACCATCACCGGTACTGGGAGATACCTCAGAGAGCAAAATCTAATGTCAAGCTCTATGTT 704  
QY 208 ValGluProValGluSerProValleuSerGlyLYsProGlyProHisLYsileGln 227  
DB 705 GTGAGCCAGTGAAGAGTCTGTTTGAATGTGGAATACTGAGACCAACAGATTCAA 764

QY 228 GlyileGlyAlaGlyPheileProGlyValleuGluValAsnleuLeuAspGluValVal 247  
DB 765 GGAATTGAGAGCTGTTTATCCCTGGAGTCTTGAGTTGATCTCCTGATGAACCTTA 824  
QY 248 GlnleSerSerAspGluAlaileGluThrAlaLYsleuAlaLeuLYsGluGlyLeu 267  
DB 825 CAGTTTCAAGTATGATGAAGCTATCGAGACTGCCAAGGCTCTGCTGAAAGAGGTTG 884  
QY 268 PheValGlyleSerSerGlyAlaAlaAlaAlaAlaPheGlnileAlaLYsArgPro 287  
DB 885 TTGTTGGAATCTCTTGTGTGAGCTGACGCTGGCGAGTTAGCTTGTAGAGGCCA 944  
QY 288 GluAlaAlaGlyLYsleuileValAlaValPheProSerPheGlyGluArgTYrleuSer 307  
DB 945 GAAACCGCCGGAAGCTATTGTTGTGTGTTCCCGAGCTTCGCGAGCGCTACCTTCA 1004  
QY 308 SerValleuPheGluSerValArgAlaGluAlaGluSerMetThrPheGluPro 325  
DB 1005 TCGTGTCTGTTCCAGTCCATCAAGAAGAGCGAGAGCATGTGTGAGGCC 1058

#### RESULT 8

US-10-425-114-23608  
Sequence 23608, Application US/10425114  
Publication No. US2004003488A1

#### GENERAL INFORMATION:

APPLICANT: Liu, Jingdong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E.  
APPLICANT: Tabaska, Jack E.  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 23608  
LENGTH: 1301  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB3597-067-G10\_FLI  
US-10-425-114-23608

#### Alignment Scores:

Pred. No.: 1.15e-149 Length: 1301  
Score: 1375.00 Matches: 268  
Percent Similarity: 91.19% Conservative: 22  
Best Local Similarity: 84.28% Mismatches: 28  
Query Match: 84.72% Indels: 0  
DB: 13 Gaps: 0

US-09-931-457a-31 (1-325) x US-10-425-114-23608 (1-1301)

QY 8 TLEAlAlAspValThrgLUleuilegLYsThrProleuValTyrlEuAsnLYsLeu 27  
DB 94 ATGCCAAGAGAGCTGACCGAGTGTATCGGACACGCCGCTGTGTACTTCAACAAGTG 153  
QY 28 AlaAspGlyCyValAlaArgValAlaAlaLYsLeuGluMetGluProCySerSer 47  
DB 154 ACCGATGGTGGTGGCGCGCGCTGCCCGCAAGCTCGAGTCCATGAGCCCTGCTCAGC 213  
QY 48 ValLYsAspArgilegLYsSerMetileAlaAspAlaGluGluLYsGlyLeuileThr 67  
DB 214 GTCAAGATAGATTGGCTACAGCATGATCAGGACGACAGAGAGAGAGGCTGATTACT 273  
QY 68 ProGlyLYsSerValleuilegLUProThrSerGlyAsnThrgLYleuAlaPhe 87  
DB 274 CCAAGCGTGAAGTGTCTGATTGAACCAACTAGCGGACACAGAGCATTTGACTGCTTT 333  
QY 88 MetAlaAlaArgGlyTyrlLYsleuileleThrMetProAlaSerMetSerLeuGlu 107



Db	334	ATGCGCTGCTGCCAAGGCGCTACAAACTTACACTCACATGCGCTGCTCCATGAGCATGGAG	393
QY	108	ArgArgIleIleLeuLeuAlaPheGlyValGluLeuValLeuThrAspProAlaGlyGly	127
Db	394	AGGAGGATCATATTGGAAGGCTTTGGTGTGACTGACTGTCTTACTGACCCACTCTGGGA	453
QY	128	MettysGlyAlaValGlnLysAlaGluGlnIleLeuAlaAlaGlyThrProAsnAlaTyrIle	147
Db	454	ATGAAAGCAGCGCTGTCAAGAAAGCGGAGAGATACAAAGCAAGACACCCAACTCGTACATT	513
QY	148	LeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThrGlyProGlu	167
Db	514	CTTCAACAATTGAAAATCCAGCTTAACCCAAAGATCACTGTGAGACTACTGGGCTTGA	573
QY	168	IleTrpLysGlySerAspGlyLysIleAspAlaPheValSerGlyIleGlyThrGlyGly	187
Db	574	ATCTGGAAGCTACACAGGAAAATTGATGGCTTGTAATCTGTATCGGACAGAGCGT	633
QY	188	ThrIleThrGlyAlaGlyLysTyrLeuLysGlnGlnAsnProAsnIleLysLeuIleGly	207
Db	634	ACCATCACCGGTACTGGCGGATACCTCAGAGAGCAAAATCCTAATGTCAAGCTCTATGCT	693
QY	208	ValGlnProValGluSerProValLeuSerGlyGlyLysProGlyProHisLysIleGln	227
Db	694	GTCGAGCCAGTGGAGAGTGTCTTTGATGTGTGAAAACCTTGACACACAGATTCAA	753
QY	228	GlyIleGlyAlaGlyPheIleProGlyValLeuGlnValAsnLeuLeuAspGluValVal	247
Db	754	GGAATTGGAGCTGCTTTTATCCCTGGAGTCTTGAGATTGATCTCTTGATGAACCTCTA	813
QY	248	GlnIleSerSerAspGluAlaIleGluThrAlaLysLeuLeuAlaLeuLysGluGlyLeu	267
Db	814	CAGGTTTCAAGTGAATGAAGCTATCGAGACTGCCAAGGCTCTTGCTGAAAGAAGGTTG	873
QY	268	PheValGlyIleSerSerGlyAlaAlaAlaAlaAlaAlaPheGlnIleAlaLysArgPro	287
Db	874	TTGGTTGGAATCTCTTCTGTGTGACAGTGCAGCTGGGCAAGTTAGGCTGTCTAAGAGGCCA	933
QY	288	GluAsnAlaGlyLysLeuIleValAlaValPheProSerPheGlyGluArgTyrLeuSer	307
Db	934	GAAACCGCCGAAAGCTATTGTGTGTGTCTTCCGAGCTTCGGCGAGCGCTACTCTCA	993
QY	308	SerValLeuPheGluSerValArgGluAlaGluSerMetThrPheGluPro	325
Db	994	TCGGTGTCTGTTCCAGTCCATCAAGAAGAAAGCGGAGAGCATGTGTGTCGAGGCC	1047

```

RESULT 9
US-09-938-842A-893
; Sequence 893, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE

```

```

? FILE REFERENCE: SCR151300-3
? CURRENT APPLICATION NUMBER: US/09/938,8
? CURRENT FILING DATE: 2001-08-24
? PRIOR APPLICATION NUMBER: US 60/227,866
? PRIOR FILING DATE: 2000-08-24
? PRIOR APPLICATION NUMBER: US 60/264,647
? PRIOR FILING DATE: 2001-01-16
? PRIOR APPLICATION NUMBER: US 60/300,111
? PRIOR FILING DATE: 2001-06-22
? NUMBER OF SEQ ID NOS: 5379
? SEQ ID NO 893
? LENGTH: 966
? TYPE: DNA
? ORGANISM: Arabidopsis thaliana
US-09-938-842A-B93

```

Alignment Scores:	
Pred. No.:	9.75e-150
Score:	1374.00
Percent Similarity:	91.08%
Best Local Similarity:	82.15%
Query Match:	84.66%
DB:	9
Length:	966
Matches:	267
Conservative:	29
Mismatches:	25
Indels:	4
Gaps:	1

US-09-931-457A-31 (1-325) X US-09-938-842A-893 (1-966)

QY	1	MetAlaValAlGluArgSerGlyTlleAlaAlaLysAspValThrGluLeuIleGlyLysThrPro	20
Db	1	ATGGCTTCAGTGGCTCCCTAAATCGCTTAAGATGTAAGTGAAGTAACTAATAGGGAACACTCCA	60
QY	21	LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu	40
Db	61	TTAGTGTATCTAAACAAAGTAGCTAAAGATTGTGTGTGTCATGTTGCTCTAAGCTTGAG	120
QY	41	LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla	60
Db	121	ATGATGGAACCATGTTCTAGCGTCAAGACAGGATTGGTTATAGTATGATTGCTGATGCA	180
QY	61	GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn	80
Db	181	GAACTTAAAGACTTATTAAACCGGAGAGAGAGTGTGTGATTGAGCCAAAGTGGGAAC	240
QY	81	ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrTyrLysLeuIleIleThrMet	100
Db	241	ACTGAGATTGGTTAGCATTTACTAGCTGACGTCGCAAGGCGTATAAGCTTGTTAAACAATG	300
QY	101	ProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluGluIleLeuAla	120
Db	301	CCAGCTTCAATGAGTATAGAGAGAGAATCATCTCTAGCTTTTGAGCTGAGTTGATT	360
QY	121	LeuThrAspProAlaLysGlyMetLysGlyAlaValGluLysAlaGluGluIleLeuAla	140
Db	361	CTTACTGATCCGGCTAAAGGTTATGTAAGCTGCTGTGCAAAAGCGGAGAGATTGGCG	420
QY	141	LysThrProAsnAlaTyrIleLeuGluGluPheGluAsnProAlaAsnProLysValHis	160
Db	421	AAACACCTPATGTATATGCTTCAACAGTTTGAGAAATCCTGCTAATCCAAAGATCCAT	480
QY	161	TyrGluThrThrGlyProGluIleTyrPlyGlySerAspGlyLysIleAspAlaPheVal	180
Db	481	TATGAGACTACTGACCTGAGATATGAGAAAGGTTCTGTGTGGAAGAGTGATGGCTTTGTT	540
QY	181	SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGluAsn	200
Db	541	TCGTGATTTGGTACTGGT-----GGTCTGGGAAGTATCTCAAGGAACAGAAC	588
QY	201	ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLys	220
Db	589	ACAACATAAAGCTGTATGTGTGACCTGTTGAAGCCCTATTCTGTCCGCTGGAAG	648
QY	221	ProGlyProHisLysIleGluGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal	240
Db	649	CCAGGTCCCCACAAATTCAGGTTAGGTGCTGCTTTTATTCAGGCATTTTCGATGTT	708
QY	241	AsnLeuLeuAspGluValValGluIleSerSerAspGluAlaIleGluThrAlaLysLeu	260
Db	709	GATCTTATAGATGAAGTTGTTCAAGTTCAAGTGAAGAATCCATTGACATGCAAGGCTT	768
QY	261	LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla	280
Db	769	CTTGCTAGGGAAGAGGCTCTGTAGTGGGAATCTCATCTGGAGCAGCAGCTACTGCAGCA	828
QY	281	PheGluIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValaPheProSer	300
Db	829	ATCAAACCTTGCAAGAAGGCCAGAAATCTCGGAAGCTTATAGTGGCGGTGTTCCGAGT	888
QY	301	PheGlyGluArgTyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSer	320



QY 3 ValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrProLeuVal 22  
 DB 277 CTCGAGGCGCTCAACATCGCCGACGACGATCAACCTCATCGGCAAGACTCCAAATGTA 336  
 QY 23 TyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMet 42  
 DB 337 TATCTCAACAACATCGTCAAGGATGTGTGCAATGTGCTGCTAAGCTCGAATTATG 396  
 QY 43 GluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGlu 62  
 DB 397 GAGCCCTGTGACGTGCAAGGACAGATAGATACAGTATGATTTCTGATGCGGAAGAG 456  
 QY 63 LysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGly 82  
 DB 457 AAAGCTTGTATACCTCTGGAAGAAGTGTGTTGGTGAACCAACAAGTGAATAACAGGC 516  
 QY 83 IleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProAla 102  
 DB 517 ATTGGTCTTCTTCATTCCTGCTCTCCAGAGATATAAATTAATATGACCAATGCTGCA 576  
 QY 103 SerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThr 122  
 DB 577 TCAATGACATGAGAGAGAGATCTTCACTCAACCTTTTGGCGCTGAACCTTGTCTTACT 636  
 QY 123 AspProAlaLysGlyMetLysGlyAlaValAlaGluLysAlaGluIleLeuAlaLysThr 142  
 DB 637 GATGCCGCAAAAGGATGAGAGGGGCTGTAGATAGGCTACAGAGATTTTAATAAGACA 696  
 QY 143 ProAsnAlaTyrIleLeuGluGlnPheGluAsnProAlaAsnProLysValHisTyrGlu 162  
 DB 697 CCTGATGCTATATGCTGACAGAGGATTCGACCAACCTGCAACCCAAAGGTAATATGAG 756  
 QY 163 ThrThrGlyProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheValSerGly 182  
 DB 757 ACTACTGGGCGAAGATCTGGAGAGATTCFAAAGGAGAGGTGATGTATTCATTGCTGGA 816  
 QY 183 IleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsnProAsn 202  
 DB 817 ATTGAACAGGTGAGACATATCTGCTGCTGCGCTTCTGAAAGAGAAATCTCTGGA 876  
 QY 203 IleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLysProGly 222  
 DB 877 ATTAAGGTATTGTATGATGACCTTCTGAGAGTACATACCTCTGCTGGAATCTGCGC 936  
 QY 223 ProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeu 242  
 DB 937 CCACATAGATTCAGGCAATGGGCGAGATTTGTTCCAAAGAACTGTGATGTGAAGTT 996  
 QY 243 LeuAspGluValAlaGlnIleSerSerAspGluAlaIleGluThrAlaLysLeuAla 262  
 DB 997 CTCGATGAGTATGATGATATCTAGTATGAGGCTGTGAGACAGCAAGCAATTTGGCT 1056  
 QY 263 LeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAlaPheGln 282  
 DB 1057 CTTCAAGGAAGATTACTGCTTGAATTTCATCTGGGGCAGACAGACAGCTGCCATTAA 1116  
 QY 283 IleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaAlaPheProSerPheGly 302  
 DB 1117 GTTCAAAAAGACCAAGAAATGCTGGAAGTGTGTGTGTGTTCCAGCTTTGGT 1176  
 QY 303 GluArgTyrLeuSerSerValLeuPheGluSerValArgGluAlaGluIleSerMetThr 322  
 DB 1177 GAGAGTACCTTTCATCTATCTTTTCACTGATAGAGAAAGATGTGAGAGTTGCAA 1236  
 QY 323 PheGluPro 325  
 DB 1237 CCTGAACCA 1245  
 RESULT 12  
 US-10-424-599-55160  
 ; Sequence 55160, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J  
 APPLICANT: Kovalic David K  
 APPLICANT: Zhou Yihua  
 APPLICANT: Cao Yongwei  
 TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(53223) B  
 CURRENT APPLICATION NUMBER: US/10/424, 599  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 285684  
 SEQ ID NO 55160  
 LENGTH: 1404  
 TYPE: DNA  
 ORGANISM: Glycine max  
 FEATURE:  
 OTHER INFORMATION: Clone ID: PAT\_MRT3847\_20820C.1  
 US-10-424-599-55160  
 Alignment Scores:  
 Pred. No.: 9.24e-143 Length: 1404  
 Score: 1316.00 Matches: 273  
 Percent Similarity: 89.67% Conservative: 22  
 Best Local Similarity: 82.98% Mismatches: 27  
 Query Match: 81.08% Indels: 8  
 DB: 13 Gaps: 3  
 US-09-931-457A-31 (1-325) x US-10-424-599-55160 (1-1404)  
 QY 1 MetalValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrPro 20  
 DB 151 ATGGCGGTGAGAGAGTGAAGCATTCGAAAGAATGTGACTGAATTCGCCGACAGACA--- 207  
 QY 21 LeuValTyrLeuAsnLysLeuAlaAspGly-CysValAlaArgVal-----AlaAl 37  
 DB 208 CTGACCTTTTG-----TTTTTGACGGAATAATTACACGATCTATCTTAACCTTGC 261  
 QY 37 aLysLeuGluLeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetI 57  
 DB 262 GAATCTAGAAATGATGGAACCTTCTCTCAAGTGTCAAGACAGAGATAGGATATAGCATGAT 321  
 QY 57 eAlaAspAlaGluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProTh 77  
 DB 322 TGTAGATGCTGAGAGAGAAAGATTCATCAACCGGTTGAAGTGTCTCATTTGAGCTTAC 381  
 QY 77 rSerGlyAsnThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuI 97  
 DB 382 TAGTGAACACAGGCGCATAGGTTTGGCATTCATGAGCACTGCTAAGGCTTATTAACCTTAT 441  
 QY 97 eIleThrMetProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAl 117  
 DB 442 TATAACCATGCTTCTTCGATGAGTCTGAGAGACAGAACCATTTCTGGGCTTTTGAGGC 501  
 QY 117 aGluLeuValLeuThrAspProAlaLysGlyMetLysGlyAlaValGluLysAlaGlu 137  
 DB 502 TGAGTTAGTCTCACTGATCCCGCAAGGTATGAAGAGAGCTGTTCACAAAGCAGAGA 561  
 QY 137 uIleLeuAlaLysThrProAsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnPr 157  
 DB 562 GATTCGGGATAGACTCCCAATTTTATATCTTCAGCAATTTGAATCTGGCAGGCC 621  
 QY 157 oLysValHisTyrGluThrThrGlyProGluIleTyrLysGlySerAspGlyLysIleAs 177  
 DB 622 AAAGATCCATATGAACAACCACTGACCGGAGATCTGAG-GGCTCTAGTGGGAAGTTGA 680  
 QY 177 PalapheValSerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLys 197  
 DB 681 TGCTCTGTCTCTGGGATTTGGAGCTGAGCTACAGTAACAGGTGTGGAATAATATCTGAA 740  
 QY 197 sGluGluAsnProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSe 217  
 DB 741 AGAGCAGATCTCATATTAAGCTATATGCAATTGAACCAAGTGAAGTCAATACTGTCTC 800  
 QY 217 rGlyGlyLysProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyVa 237



```
DB 801 TGGAGGAAAACCCGGCCCTCATAGATCAAGAAATTGGTGGCTTCATCCCTGGTGT
QY 237 lueuGluValAsnLeuAspGluValGlnIleSerSerAspGluValIleGluTh 257
DB 861 TCTGATGTTGATTACTTGAATGAAGTTGTTCAATTCAAGTGAAGAAGCTATTGAAC 920
QY 257 rAlaIleuLeuAlaIleuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaI 277
DB 921 TGCTAAGCTTCTTGCAATTGAAAGAGGTTTACTGCTGGAAATATCACTGCTGCTGC 980
QY 277 aAlaAlaIleuGlnIleAlaIleuAspProGluAsnAlaGlyLysLeuIleValAlaVa 297
DB 981 AGCTGCTGCAGTTAAGATAGCGAAGAGCCAGAAATGCTGAAACTCATTTGTCGGT 1040
QY 297 lPheProSerPheGlyGluArgTyrIleuSerSerValIleuPheGluSerValArgArgG 317
DB 1041 CTTCCCAAGTTTGAGAGAGCGGTATCTATGCTGCTCTTGAATCCGTGAAGCGAGA 1100
QY 317 uAlaGluSerMetThrPheGluPro 325
DB 1101 AGCAGAGAGCTTGTGCTTGAAGCCT 1125

RESULT 13
US-10-425-114-22525
; Sequence 22525, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 22525
; LENGTH: 1279
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3587-229-H7_FLI
US-10-425-114-22525
```

```
Alignment Scores:
Pred. No.: 5.25e-142 Length: 1279
Score: 1309.00 Matches: 258
Percent Similarity: 88.05% Conservative: 22
Best Local Similarity: 81.13% Mismatches: 26
Query Match: 80.65% Indels: 12
DB: 13 Gaps: 1
```

```
US-09-931-457A-31 (1-325) x US-10-425-114-22525 (1-1279)
QY 8 lIleAlaIleuAspValThrGluLeuIleGlyLysThrProLeuValTyrLeuAsnLysLeu 27
DB 108 ATCGCCAGAGAGCTCAACCGAG-----GTG 131
QY 28 AlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluProCysSerSer 47
DB 132 ACCGATGGGTGCGTCGGCCGCTCGCCCAAGCTCGAGTCCATGAGAGCCCTGCTCAGC 191
QY 48 ValIleAspArgIleGlyTyrSerMetIleAlaAspAlaGluGluLysGlyLeuIleThr 67
DB 192 GTCAAGGATAGGATTGGCTACAGCATATCAACGACGACGAGAGAAAGGCTGATTACT 251
QY 68 ProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIleGlyLeuAlaPhe 87
DB 252 CCAAGCGGTGAGTGTCTGATTGAACCACTAGCGGCAACAGGCAATTGAGCTGCGCTTT 311
```

```
QY 88 MetAlaAlaIleArgGlyTyrLysLeuIleIleThrMetProAlaSerMetSerLeuGlu 107
DB 312 ATGGCTGCTGCCAAGGCTTACAACTTACACTACACTGCTGCTCCCTCCATGAGCATGAG 371
QY 108 ArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspProAlaLysGly 127
DB 372 AGAGGATCATATTGAAGCTTTGGTGGTGAACCTTGCTTACTGACCCACTGCTGGGA 431
QY 128 MetLysGlyAlaValGlnLysAlaGluGluIleLeuAlaLysThrProAsnAlaTyrIle 147
DB 432 ATGAAAGAGCTGTCAAGAAAGCGGAGAGATACAGCAAGACACCAACTGCTGATC 491
QY 148 LeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThrGlyProGlu 167
DB 492 CTTCAACAATTGAAATCCAGCTAACCCAAAGATTCATATGAGACTACTGGGCTGAA 551
QY 168 lIleTyrLysGlySerAspGlyLysIleAspAlaPheValSerGlyIleGlyThrGlyGly 187
DB 552 ATCTGAAAGCTACAGCAGAAATAATGCTGCTGCTTGTATCTGTATCGGACAGAGGT 611
QY 188 ThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsnProAsnIleLysLeuIleGly 207
DB 612 ACCATCACCGGTACTGCGCGATACCTCAGAGAGCAAAATCCTAATGTCAAGCTTATGGT 671
QY 208 ValGluProValGluSerProValLeuSerGlyLysProGlyProHisTyrIleGln 227
DB 672 GTGAGCCAGTGAGAGTGTCTGTTTGAATGGTGAAACCTGACACACAGATTTCAA 731
QY 228 GlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeuLeuAspGluValVal 247
DB 732 GGAATTGAGCTGTTTATCCCTGAGTCTTGATGTTGATCTCCTGATGAGACTCTA 791
QY 248 GlnIleSerSerAspGluAlaIleGluThrAlaLysLeuAlaLeuLysGluGlyLeu 267
DB 792 CAGGTTCAAGTATGATGATGATGAGACTGCCAGGCTTGCTCTGAAGAAGGTTG 851
QY 268 PheValGlyIleSerSerGlyAlaAlaAlaAlaAlaPheGlnIleAlaLysArgPro 287
DB 852 TTGGTTGATCTCTTCTGCTGAGCTGCACTGCGGCAAGTAAAGCTTGCTGAAGAGGCCA 911
QY 288 GluAsnAlaGlyLysLeuIleValAlaValPheProSerPheGlyGluArgTyrLeuSer 307
DB 912 GAAACGCCGGAAGCTATTGTTGTGCTGTTCCGAGCTTCGGCGAGCGCTACTCTCA 971
QY 308 SerValLeuPheGluSerValArgArgGluAlaGluSerMetThrPheGluPro 325
DB 972 TCGGTGCTGTTCCAGTTCATCAAGAGAGAGCGGAAAGCATGTGTGAGGCC 1025
```

```
RESULT 14
US-10-425-114-34372
; Sequence 34372, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 34372
; LENGTH: 1490
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLM017258D02_FLI
US-10-425-114-34372
```

Alignment Scores:

Pred. No.:	2,01e-138	length:	1490
Score:	1279.00	Matches:	245
Percent Similarity:	88.38%	Conservative:	44
Best Local Similarity:	74.92%	Mismatches:	36
Query Match:	78.80%	Indels:	2
DB:	13	Gaps:	1

US-09-931-457A-31 (1-325) X US-10-425-114-34372 (1-1490)

Qy	1	MetalAlaValGluArgSergly-----IleAlaLysAspValThrGluLeuIleGlyLys	18
Db	213	GTGCGCCGCGAGGTCGGGGGCGCTCAACATCGCCAAAGACGCTCAACCCAGCTCATCGGCAAC	272
Qy	19	ThrProLeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLys	38
Db	273	ACACCAATGGTGTAATCTCAACAACGTCGTCAGGGATCTGTGGCAATGTCGCTTAAG	332
Qy	39	LeuGluLeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAla	58
Db	333	CTCGAGATTATGGAGCCCTGCTGTAGCGTGAAGGACAGGATAGGGTACAGTATGATAAAT	392
Qy	59	AspAlaGluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSer	78
Db	393	GATGCTGAACAGAAAGGCGTTGATTACTCTCGAAGAAGTGTTTGGTGAAGCAACAACT	452
Qy	79	GlyAsnThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIle	98
Db	453	GGAACACACAGGCAATTGGTCTTGCCCTTCATTGCTGCTCTTAAGAGATATAAGCTGATACTA	512
Qy	99	ThrMetProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGlu	118
Db	513	ACAATGCCCTTCATCAATGAGCATGAGAGAGAGAGTCTCTCTCAGAGCTTTTGGTCTGAA	572
Qy	119	LeuValLeuThrAspProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluGluIle	138
Db	573	CTTGCTCTTACTGATGCTGCAAAAGGAGTGAAGGGCGCTTAGATTAAGGCTTACAGAGATT	632
Qy	139	LeuAlaLysThrProAsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLys	158
Db	633	TTAAACAAGACACCAAAATTCCTACATGCTTCAACAGTTGATTAACCTTGCCAACTTAAG	692
Qy	159	ValHisTyrGluThrThrGlyProGluIleTribLysGlySerAspGlyLysIleAspAla	178
Db	693	GTACATTATGAGACTACTGCTCCAGAGATCTGGAGAGATTCAAGGGGAAGGTGATATA	752
Qy	179	PheValSerglyIleGlyThrGlyGlyTyrIleThrGlyAlaGlyLysTyrLeuLysGlu	198
Db	753	TTCATTGGTGAATTTGAACAGGGGAGACAAATCTGTGTGCGCGCGTTTCTCAAGAG	812
Qy	199	GlnAsnProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSergly	218
Db	813	AAAAATCCCGAATTAAAGTTATGTATTTGAAGCTTCTGAAAGTAACATACTCTCCGGT	872
Qy	219	GlyLysProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeu	238
Db	873	GGAAGAACCCTGGTCCACATAGATCCAGGAATCGGCGCAGGATTGTTCAGAAGACTTG	932
Qy	239	GluValAsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAla	258
Db	933	GATAGCGATATCTTGATGAAGTAATTGAGATATCAAGTATGTAAGCTGTGAAGACGA	992
Qy	259	LysLeuLeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerglyAlaAlaAlaAla	278
Db	993	AAACAGTTGGCTGTTCAGGAAGGATTACTGTGGAATCTCTCTGCGGCGAGCGCGCC	1052
Qy	279	AlaAlaPheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPhe	298
Db	1053	GCCGCATTAAGGTTGCCAAAAGACCAAGAGATGCTGAAAAGCTGATAGTGTGCTGTTT	1113
Qy	299	ProSerPheGlyGluArgTyrLeuSerSerValLeuPheGluSerValArgGluAla	318

Db	1113	CCGAGCTTCGGCGAGAGTAACCTTTCATCTGTCTCTATCAGTCCATAAGAGAATGC	1172
QY	319	GlutMetThrPheGluPro	325
Db	1173	GAGAACATGCAGCCTGAGCCA	1193
RESULT 15			
US-10-425-114-34527			
/ Sequence 34527, Application US/10425114			
/ Publication No. US20040034888A1			
/ GENERAL INFORMATION:			
/ APPLICANT: Liu, Jindong			
/ APPLICANT: Zhou, Yihua			
/ APPLICANT: Kovalic, David K.			
/ APPLICANT: Screen, Steven E			
/ APPLICANT: Tabaska, Jack E			
/ APPLICANT: Cao, Yongwei			
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With			
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement			
/ FILE REFERENCE: 38-21(53313)B			
/ CURRENT APPLICATION NUMBER: US/10/425,114			
/ CURRENT FILING DATE: 2003-04-28			
/ NUMBER OF SEQ ID NOS: 73128			
/ SEQ ID NO 34527			
/ LENGTH: 1491			
/ TYPE: DNA			
/ ORGANISM: Zea mays			
/ FEATURE:			
/ OTHER INFORMATION: Clone ID: UC-ZMFLMO17281G01_FLI			
US-10-425-114-34527			
Alignment Scores:			
Pred. No.: 2.01e-138 Length: 1491			
Score: 1279.00 Matches: 245			
Percent Similarity: 88.38% Conservative: 44			
Best Local Similarity: 74.92% Mismatches: 36			
Query Match: 78.80% Indels: 2			
DB: 13 Gaps: 1			
US-09-931-457A-31 (1-325) x US-10-425-114-34527 (1-1491)			
QY	1	MetaIaValGIuArgSerGly-----IleAlaIysAspValThrGluLeuIleGIlyLys	18
Db	214	GTGCGCGCCGAGTCGGGGGCTCAACATCGCCACGACGACGACCCAGCTCATCGGCAAC	273
QY	19	ThrProLeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaIalys	38
Db	274	ACACCAATGGTGTATCTCAACAACGTGCTCAAGGATCTGTGCGCAATGTCGCTGAAG	333
QY	39	LeuGluLeuMetGluProCysSerSerValLysAspArgIleGIlyTyrSerMetIleAla	58
Db	334	CTCGAGATTATGAGAGCCCTGCTGTAGCGTGAAGGACAGATAGGTTACGATTAAT	393
QY	59	AspAlaGluGluLysGlyLeuIleThrProGIlyLysSerValLeuIleGluProThrSer	78
Db	394	GATGCTGAACAGAGGCGCTGATTACTCTCTGAAAGAGTGTTTGGTGAAGCAACAAGT	453
QY	79	GlyAsnThrGlyIleGIlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIle	98
Db	454	GGAACACAGCGCATTTGCTCTTCCTTCATTGCTGCTCTTAAGAATATAAGCTGATACTA	513
QY	99	ThrMetProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGlu	118
Db	514	ACAATGCTTCATCATGACATGAGATGAGAGAGAGTCTCTCAGAGCTTTTGGTGTGAA	573
QY	119	LeuValLeuThrAspProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluGluIle	138
Db	574	CTTGCTCTTACTGATGCTGCAAAAGGATGAAGGGGCTTAGATAAGGCTACAGAGATT	633
QY	139	LeuAlaIysThrProAsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLys	158
Db	634	TTAAACACAGACCAATTTCTTACATGCTTCAACAGTTGATTAACCTCGCAACCTAAG	693

QY 159 ValHisTyrGluThrThrGlyProGluIleTryptylSerAspGlyLysIleAspAla 178  
 Db 694 GTACATTATGAGACTGCTGCTCCAGAGATCTGGGAGATTCAAAGGGAGGTGATATA 753  
 QY 179 pheValSerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGlu 198  
 Db 754 TTCATTGGTGAATTGGAACAAGGGGAGCAATATCTGTGCCGCGTTTCTCAAGAG 813  
 QY 199 GluAsnProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGly 218  
 Db 814 AAAATCCTGGAATTAGCTTATGTAATGAGCTTCTGAAGTAACATACCTCTCCGGT 873  
 QY 219 GlyLysProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeu 238  
 Db 874 GGAATAACCTGTGCTCCACATAGATCCAGGGAATCGGCGAGGATTTGTTCCAGCACTTG 933  
 QY 239 GluValAsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAla 258  
 Db 934 GATAGCGATATTCTTGATGAAGTAATTGAGATATCACTGATGAAGCTGTGAGACAGCA 993  
 QY 259 LysLeuLeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAla 278  
 Db 994 AAACAGTTGCTGCTCAGGAAGGATTAATCTGCTGGAATCTCTCTGCGGCGAGCCGCC 1053  
 QY 279 AlaAlaPheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPhe 298  
 Db 1054 GCCGCCATTAAAGGTTGCCAAAGACCAAGAAATGCTGGAAGCTGATAGTGTGCTGTT 1113  
 QY 299 ProSerPheGlyGluArgTyrLeuSerSerValLeuPheGluSerValArgArgGluAla 318  
 Db 1114 CCGAGCTTCGGCGAGAGTAACCTTCATCTGTCTCTATCAATCCATAGAGAGAAATGC 1173  
 QY 319 GluSerMetThrPheGluPro 325  
 Db 1174 GAGAACATGACGCTGAGCCA 1194

Search completed: May 22, 2004, 01:56:27  
 Job time : 588 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2004, 14:15:53 ; Search time 5582 Seconds  
(without alignments)  
10575.638 Million cell updates/sec

Title: US-09-931-457A-30  
Perfect score: 1362  
Sequence: 1 acttgtagtcgtagatag.....aaaaaaaaaaaaaaaa 1362

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_scs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_scs:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rtd:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1238	90.9	1267	8	AF452451	AF452451 Glycine m
2	707.4	51.9	953	8	CAR6024	AU06024 Cicer ari
3	656.4	48.2	1247	8	AY450295	AY450295 Nicotiana
4	645	47.4	1855	8	CNAPCCS7	D28777 Citrullus v
5	643.2	47.2	1354	8	AB029511	AB029511 Solanum t
6	642.4	47.2	1253	8	ATOACLY	X80376 Arabidopsis
7	640.8	47.0	1332	8	ATCYS3A	X84097 A.thaliana
8	636.4	46.7	1280	8	BJY10845	Y10845 Brassica ju
9	636.4	46.7	1281	8	SPIOL	D10476 Spinach mRN
10	636.4	46.7	1303	6	E06716	B06716 CDNA encodi
11	632.8	46.5	1308	8	AF044172	AF044172 Solanum t
12	632.6	46.4	1233	8	AY063827	AY063827 Arabidops
13	628	46.1	1254	8	AY045825	AY045825 Arabidops
14	618	45.4	1238	8	BJY10847	Y10847 Brassica ju
15	609.4	44.7	1192	8	AY353092	AY353092 Populus x
16	595.2	43.7	966	6	AX506198	AX506198 Sequence
17	595.2	43.7	966	6	AX651825	AX651825 Sequence
18	586.4	43.1	1671	8	AB040503	AB040503 Allium tu
19	557.8	40.6	1378	8	AB029512	AB029512 Solanum t
20	552.6	40.6	1319	8	ZMCSOATL	X85803 Z.mays mRNA
21	551.4	40.5	1404	8	AF044173	AF044173 Solanum t
22	545.2	40.0	993	6	AX652865	AX652865 Sequence
23	545.2	40.0	1280	8	AK061993	AK061993 Oryza sat
24	545.2	40.0	1281	8	AK099598	AK099598 Oryza sat
25	543.6	39.9	1290	8	AF073695	AF073695 Oryza sat
26	541.4	39.8	1368	8	ATH011976	AJ011976 Arabidops
27	538.4	39.5	1363	8	AF073697	AF073697 Oryza sat
28	538.4	38.9	1355	8	BT009041	BT009041 Triticum
29	530.4	38.9	5259	6	AK653890	AK653890 Sequence
30	529.4	38.9	1371	8	WHICYS1	D13153 Triticum ae
31	528.8	38.8	1278	8	ATCTACSI	X81697 A.thaliana
32	525.6	38.6	1483	6	E08016	E08016 CDNA encodi
33	524.6	38.5	1484	8	SPICYSY	D14722 Spinacia ol
34	524.6	38.5	1484	8	SPICYSY	AK065007 Oryza sat
35	522.2	38.3	2104	8	AK065652	AK065652 Oryza sat
36	522.2	38.3	1562	8	AK099217	AK099217 Oryza sat
37	521.6	38.2	1555	8	AK104704	AK104704 Oryza sat
38	520	38.2	1432	8	SOCYSK	X66860 S.oleracea
39	518.2	37.9	1493	8	NTA299249	AJ299249 Nicotiana
40	516.2	37.6	1377	8	ATH271727	AJ271727 Arabidops
41	511.8	37.6	1460	8	AY128885	AY128885 Arabidops
42	511.8	37.6	1556	8	AY099721	AY099721 Arabidops
43	509.6	37.4	1210	8	AY096681	AY096681 Arabidops
44	509.6	37.4	1378	8	ATOACLY2	X80377 Arabidopsis
45	509.6	37.4	1378	8	ATOACLY2	X80377 Arabidopsis

ALIGNMENTS

RESULT 1  
AF452451 1267 bp mRNA linear PLN 21-JAN-2002  
LOCUS AF452451  
DEFINITION Glycine max cysteine synthase mRNA, complete cds.  
ACCESSION AF452451  
VERSION AF452451.1 GI:18252505  
KEYWORDS  
SOURCE  
ORGANISM Glycine max (soybean)  
Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
REFERENCE 1 (bases 1 to 1267)  
AUTHORS Demosthenis,C. and Krishnan,H.B.

TITLE Nucleotide sequence of a soybean cDNA encoding cysteine synthase  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1267)  
AUTHORS Demosthenis, C. and Krishnan, H.B.  
TITLE Direct Submission  
JOURNAL Submitted (28-NOV-2001) USDA-ARS, University of Missouri, 108W  
Curtis Hall, Columbia, MO 65211, USA

FEATURES  
Source  
1. 1267  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
82. 1059  
/codon\_start=1  
/product="cysteine synthase"  
/protein\_id="AA166291.1"  
/translation="MAVERSGIAKDYTELIGKPLVLYLNKLDGCVARVAKLILMEP  
CSSVKDRIGSMIADAEKGLITPKSVLIBPTSGNTGIGLAEMAAARGYKLITMPA  
SMSIERIILAFGAELVITDPKMGKAVOKABILAKTPNAVILQFENPANKVH  
YETGPRIKSGSDKIDAFVSGITGTTGAGKYLKEQNPITKLIGVBPVPSVLSG  
GXPHKIOGIGAGFIPGVLEVNILDEVLOISSDBAIEITAKLALKEGLFVGISSGAA  
AAAAFOIAKRPENAKLIVAVFSPGFRYLSSVLFEVRRBAESMTFEP"

ORIGIN  
Query Match 90.9%; Score 1238; DB 8; Length 1267;  
Best Local Similarity 99.6%; Pred. No. 5.9e-291;  
Matches 1241; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 20 GCCGATGCTGCTGTAGTGTCTCACTCATTCCTGTTCTCAAGTCAAGCTTTAGTGTG 79  
DB 12 GCCGATGCTGCTGTAGTGTCTCACTCATTCCTGTTCTCAAGTCAAGCTTTAGTGTG 71  
QY 80 AGCAGATATATGCTGTGAAAGGTCCGGAATGCCAAGATGTACGGAATTGATTGG 139  
DB 72 AGCAGATATATGCTGTGAAAGGTCCGGAATGCCAAGATGTACGGAATTGATTGG 131  
QY 140 TAAACCCCATTTAGTATCTAATAAATCTGCGGATGCTGTGTTGCCCGGTTGCTGC 199  
DB 132 TAAACCCCATTTAGTATCTAATAAATCTGCGGATGCTGTGTTGCCCGGTTGCTGC 191  
QY 200 TAACTGAGTTGATGAGCCATGCTCTAGTGTGAAGACAGAGATTGGGTATGATGAT 259  
DB 192 TAACTGAGTTGATGAGCCATGCTCTAGTGTGAAGACAGAGATTGGGTATGATGAT 251  
QY 260 TGCTGATGAGAGAGAGGGAATTATCACACCTGGAAAGAGTGTCTCATTTAGCCCAAC 319  
DB 252 TGCTGATGAGAGAGAGGGAATTATCACACCTGGAAAGAGTGTCTCATTTAGCCCAAC 311  
QY 320 AAGTGTATATCTGCAATGATTTAGCTTCATGCGACAGCCAGGGGTTACAAAGCTCAT 379  
DB 312 AAGTGTATATCTGCAATGATTTAGCTTCATGCGACAGCCAGGGGTTACAAAGCTCAT 371  
QY 380 AATTACAATGCTGCTTCTATGAGTCTTGAGAGAAATCATTTCTATTAGCTTTTGAGC 439  
DB 372 AATTACAATGCTGCTTCTATGAGTCTTGAGAGAAATCATTTCTATTAGCTTTTGAGC 431  
QY 440 TGAGTGTGTTGACAGATCCTGCTAAGGGAATGAAGGTGCTGTTCAAGAGGCTGAAGA 499  
DB 432 TGAGTGTGTTGACAGATCCTGCTAAGGGAATGAAGGTGCTGTTCAAGAGGCTGAAGA 491  
QY 500 GATATTGGCTAAGAGCCCAATGCTACATCTTCAACATTTGAAAACTCTGCCAATCC 559  
DB 492 GATATTGGCTAAGAGCCCAATGCTACATCTTCAACATTTGAAAACTCTGCCAATCC 551  
QY 560 CAAGGTTCAATATGAAACCACTGCTCAGAGATATGAAAGGCTCCGATGGGAAAAATTGA 619  
DB 552 CAAGGTTCAATATGAAACCACTGCTCAGAGATATGAAAGGCTCCGATGGGAAAAATTGA 611  
QY 620 TGCATTTGTTCTGGGATAGGCACTGCTGATCAATAACAGGTGCTGGAATAATCTTAA 679  
DB 612 TGCATTTGTTCTGGGATAGGCACTGCTGATCAATAACAGGTGCTGGAATAATCTTAA 671

QY 680 AGAGCAATCCGATATATAAGCTGATGCTGTGGAACCACTTGAAGTCCAGTCTCTC 739  
DB 672 AGAGCAATCCGATATATAAGCTGATGCTGTGGAACCACTTGAAGTCCAGTCTCTC 731  
QY 740 AGAGGAAAGCTGTGTCACAGAGATTCAGGAGATTGGTGTGTTTATCCCTGTGT 799  
DB 732 AGAGGAAAGCTGTGTCACAGAGATTCAGGAGATTGGTGTGTTTATCCCTGTGT 791  
QY 800 CTGGAAGTCAATCTTCTGATGAGTGTCTCAATATCAAGTGAAGCAATAGAAAC 859  
DB 792 CTGGAAGTCAATCTTCTGATGAGTGTCTCAATATCAAGTGAAGCAATAGAAAC 851  
QY 860 TGCAAGCTCTTGGCGTTAAGAGGCTATTTGTGGATATCTCCGAGCTGACG 919  
DB 852 TGCAAGCTCTTGGCGTTAAGAGGCTATTTGTGGATATCTCCGAGCTGACG 911  
QY 920 TGCTGCTGCTTTTCAAGATTGCAAAAAGACCAAGAAATGCCGGAAGCTTATTTGCCGT 979  
DB 912 TGCTGCTGCTTTTCAAGATTGCAAAAAGACCAAGAAATGCCGGAAGCTTATTTGCCGT 971  
QY 980 TTTTCCGACTTCGGGAGAGGTAACCTGCTCTCCGCTGATTTTGAAGTCAAGTGAAGCGCA 1039  
DB 972 TTTTCCGACTTCGGGAGAGGTAACCTGCTCTCCGCTGATTTTGAAGTCAAGTGAAGCGCA 1031  
QY 1040 AGCTGAAAGCATGACTTTTGAGCCCTGAATTCCTCCGTTTAAAGCTCTCACTAGATTTT 1099  
DB 1032 AGCTGAAAGCATGACTTTTGAGCCCTGAATTCCTCCGTTTAAAGCTCTCACTAGATTTT 1091  
QY 1100 CTGTGTAATGTAACAGGCTTTAAGTGTGTTAGAGTACTGTTGTGACTCTGAC 1159  
DB 1092 CTGTGTAATGTAACAGGCTTTAAGTGTGTTAGAGTACTGTTGTGACTCTGAC 1151  
QY 1160 TCTAAATATAAAGCTGCTCCAAAGACTAGTTTCTTGATGCCCCCTGAGCGATAATT 1219  
DB 1152 TCTAAATATAAAGCTGCTCCAAAGACTAGTTTCTTGATGCCCCCTGAGCGATAATT 1211  
QY 1220 TGTGCTGCAACATTAAAAAGTATCAAAAGTGTCTTATTAAGTACA 1265  
DB 1212 TGTGCTGCAACATTAAAAAGTATCAAAAGTGTCTTATTAAGTACA 1257

RESULT 2  
CAR6024 953 bp mRNA linear PLN 11-MAY-1998  
LOCUS Cicer arietinum mRNA for cysteine synthase, partial.  
DEFINITION  
ACCESSION AJ006024  
VERSION AJ006024.1 GI:3127889  
KEYWORDS cysteine synthase.  
SOURCE Cicer arietinum (chickpea)  
ORGANISM Cicer arietinum  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Ciceraceae;  
Cicer.  
REFERENCE 1  
AUTHORS Dopico, B., Esteban, R. and Labrador, E.  
TITLE cDNA expressed in chickpea epicotyls  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 953)  
AUTHORS Labrador, E.  
TITLE Direct Submission  
JOURNAL Submitted (04-MAY-1998) Labrador E., Dpto. Fisiologia Vegetal,  
Univ. Salamanca, Campus Miguel de Unamuno, Pza. Doctores de la  
Reina s/n, 37007-Salamanca, SPAIN  
FEATURES  
Source  
1. 953  
/organism="Cicer arietinum"  
/mol\_type="mRNA"  
/cultivar="Castellana"  
/db\_xref="taxon:3827"  
/clone="Cancys"  
/tissue\_type="etiolated epicotyls"  
/dev\_stage="5 day old seedlings"

```

CDS
    <1. .803
    /EC_number="4.2.99.8"
    /codon_start=3
    /product="cysteine synthase, O-acetyl-L-serine
    {thiol}-lyase"
    /protein_id="CAA06819.1"
    /db_xref="GI:3127890"
    /db_xref="GOA:O65747"
    /translation="ASPEKGLITPGQSVLIEPTSGNTIGLAFMAAKGYKLIITPAS
    MSLERRIILAFGAEVLVLTDPAKEMKGAVGRAEILAKTPNAYILQEPENPANKVHY
    ETTGPEIKMGTEGKVDAPVSGITGTITGAGKYLKQNPNIKLGVGEPVSSVLSGGAA
    KGPFRKIQIGAGFVPGVLEVNLIDEVIISSDEAIEIAKLALKEGLFVGISSGAAAA
    AAKIARIKRPENAGKLIIVVFPSPFGERYLLSVLFESVRRQAETMTFEP"

```

Query Match	51.9%	Score 707.4	DB 8	Length 953
Best Local Similarity	91.9%	Pred. No. 1,4e-161		
Matches 747	Conservative 0	Mismatches 66	Indels 0	Gaps 0
QY 265	ATGCAGAGAGAGGGAAGGCTTATTCACACCTGGAAAGAGTGCTCTATTGAGCCACAACTG	324		
DB 1	ATGCAGAGAGAGGAGGGCTTATTCACACCTGGACAGAGTGCTCTATTGAAACCAACAACTG	60		
QY 325	GTAACTACTGSCATTGATTAGCCTTCATGGCAGCAGCGGGGTTACAGCTCATTAATTA	384		
DB 61	GTAACTACTGSCATTGATTAGCCTTATATGCGAGCGGCAAGGGTTACAGCTCATTAATTA	120		
QY 385	CAATGCCCTGCTTCTATGAGCTTGAGAGAGAAATCATTTCTATTAGCTTTTGAAGCTAGT	444		
DB 121	CGATGCCCTGCTTCTATGAGCTTGAAAGAGAAATCATTTCTATTGCTTTTGAAGCTAGT	180		
QY 445	TGTTTCTGACAGATCCTGCTAAGGGAATGAAGGTGCTGTTCAAGAGCTGAAGAGATAT	504		
DB 181	TGTTTCTGACAGATCCTGCTAAGGGAATGAAGGTGCTGTTCAAGAGCTGAAGAGATAT	240		
QY 505	TGGCTAGACGCGCCATGCTTACATATCTTCAACAATTTGAAGAACCTGCCAATCCCAAG	564		
DB 241	TGGCTAGACACCCATGCTTACATATCTTCAACAATTTGAAGAACCTGCCAATCCCAAG	300		
QY 565	TTCAATTATGAACCACTGCTCCAGAGATATGGAAGGCTCCGATGGGAAATTGATGAT	624		
DB 301	TACATTATGAACCACTGCTCCAGAGATATGGAAGGCAAGAGGAAAGTTGACGCAAT	360		
QY 625	TTGTTTCTGGGATAGGCACTGCTGTACATTAACAGTGCTGGAATATCTTAAGAGC	684		
DB 361	TTGTTTCTGGGATAGGCACTGCTGTACATTAACAGTGCTGGAATATCTTAAGAGC	420		
QY 685	AGAATCCGAATATAAGCTGATTTGGTGTGGAACCAAGTTGAAAGTCCAGTGCTTCAGAG	744		
DB 421	AAATCCCAATATAAGCTGATTTGGTGTGGAACCGGTGGAAGTCCAGTGCTTCAGAGG	480		
QY 745	GAAAGCCTGTCCACACAAGATTCAAGGGAATGGTGTGTTTATCCCTGTGCTTGG	804		
DB 481	GAAAGCCTGTCTCCACAAGATTCAAGGGAATGGTGTGTTTATCCCTGTGCTTGG	540		
QY 805	AAGTCAATCTTCTGATGAAGTTGTTCAATATCAAGTGAAGAGCAATAGAACTGCAA	864		
DB 541	AAGTCAATCTTATGATGAAGTCAATCAATATCAAGTGAAGAGCAATAGAACTGCAA	600		
QY 865	AGCTTCTTGCGCTTAAAGAGGCTATTTGTGGAATATCTTCCGAGCTGCAGTCTG	924		
DB 601	AGCTTCTTGCGCTTAAAGAGGCTTTTGTGGAATATCTTCCGAGCTGCAGGCGG	660		
QY 925	CTGCTTTTCAAGTTGCAAAAAGACCAAGAAATGCCGGGAAGCTTATTTGTGCGGTTTTT	984		
DB 661	CTGCCATTAAAGTACCCAAAAGGCCAGAAATGCTGGGAAGCTTATTTGTGCTTTTTT	720		
QY 985	CCAGCTTCCGGGAGAGGTACCTGTCTCCGTGCTATTGAGTCAAGTGAAGCGCGAAGCTG	1044		
DB 721	CAAGCTTTGGGAGAGGTACTTGTCTCTGTGCTATTGAGTCAAGTGAAGCGTCAAGCTG	780		
QY 1045	AAAGCATGACTTTTGAAGCCCTGAATTTCCCGTTT	1077		

Db 781 AAACCATGACTTTCGAGCCCTGAATTATTGTT 813

RESULT 3	AY450295	1247 bp	mRNA	linear	PLN 22-NOV-2003
LOCUS	AY450295				
DEFINITION	Nicotiana glumbaginifolia cysteine synthase (Csl) mRNA, partial cds.				
ACCESSION	AY450295				
VERSION	AY450295.1	GI:38350578			
KEYWORDS					
SOURCE	Nicotiana plumbaginifolia (curled-leaved tobacco)				
ORGANISM	Nicotiana plumbaginifolia				

REFERENCE	1 (bases 1 to 1247)
AUTHORS	Liszewska, F. and Sirko, A.
TITLE	Isolation of Nicotiana glumbaginifolia cDNAs encoding O-acetylserine (thiol) lyase and serine acetyltransferase
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 1247)
AUTHORS	Liszewska, F. and Sirko, A.
TITLE	Direct Submision
JOURNAL	Submitted (26-OCT-2003) Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Pawinskiego 5A, Warsaw 02-106, Poland
FEATURES	Location/Qualifiers

```

/organism="Nicotiana glauca"
/mol_type="mRNA"
/db_xref="taxon:4092"
<1..1247
/gene="Csl"
<1..972
CDS

```

```

/EC_number="2.5.1.47"
/note="NPCSL; O-acetylserine (thiol) lyase"
/codon_start=1
/product="cysteine synthase"
/protein_id="AAR18402.1"
/db_xref="GI:38350579"
/translation="GEKNGIAKDYTELLIGNTPLVYINNVVDGCVARVAAXIESMEPCS
SVKDRIGYSMTDABEKGLIKPGSVLIEPTSGNTGVGLAFMAARGKLIITPMSSM
SLERRILRAFGAEVLVTDPAKMGKSIQAKAEIKATPNSFILOFPNPANPKVHYE
TTGPEIWKSGNKVDALVSGIGTSGITIGSGKLRQNPDIKLVGPEVESAIIISGK
PGPHKIOGIGAGFIPGVLEGLIDETIQQSSERAIETAKLTLAKGLLWGISGASAA
AAITLAKRPENAGKLIVVFPSEGEERYLISSVLFBSVRRREANNVTEP"

```

Query Match	48.2%;	Score 656.4;	DB 8;	Length 1247;
Best Local Similarity	79.8%;	Pred. NO. 3.6e-149;		
Matches 774;	Conservative 0;	Mismatches 196;	Indels 0;	Gaps 0;

QY	99	GAAAGSTCCGGAAATTGCCAAGAGTGTACGGAAATTGATTGGTAAACCCCATTTAGTATAT	158
Db	4	GAAAAGATGGAATTGCCAAGAGTGTAACTGAATTGATCGGTACACTCTCTTTGGTAC	63
QY	159	CTAAATTAACCTTCGGATGGTGTGTGTTCCCGGGTGTCTGCTAAACTGAGCTTGATGGAG	218
Db	64	CTGAATTAATGTTGTGATGGGTGTGTTCCCGCGTGTCTGCCAAGCTTGAAAGCATGGAG	123
QY	219	CCATGCTCTAGTGTGAAGCAGACAGATTGGGTATAGTATGATTGCTGATGCAGAGAGAAG	278
Db	124	CCATGCTCTAGTGTGAAGATAGATTGGTTATAGTATGATTACAGATGCTGAGAGAGAAA	183
QY	279	GGACTTATCAACCTGGAAGAGTGTCTCATTTAGCCCAACAAGTGTAACTAGCTGCATT	338
Db	184	GGCTGTATCAAACTGGCGAGAGTGTCTCATTTAGAACTACAAAGTGAAACACTGGAGTA	243
QY	339	GGATTAGCCTTCATGGCAGCAGCCAGGGGTACAGAGTCAATAATTACAAATGCCCTTCT	398
Db	244	GGATTGGCAATTATGGCTGTCTGCTAAAGGCTACAAACTCATCATTAAGATGCTTCTTCA	303



QY 399 ATGAGTCTTGAGAGAGAAATCATTCTATTAGCTTTGGAGCTGAGTTGGTTCTGACAGAT 458  
 DB 304 ATGAGTCTTGAGAGAGAAATCATTCTATTAGCTTTGGAGCTGAGTTGGTTCTGACAGAT 363  
 QY 459 CCTGCTAAGGGAATGAAGAGTCTGTCAGAGAGCTGAGAGATATTTGGCTAAGAGCGCC 518  
 DB 364 CCAGCAAAAGGGAATGAAGAGTCTGTCAGAGAGCTGAGAGATATTTGGCTAAGAGCGCC 423  
 QY 519 AATGCTACATCTTCAACAATTTGAAAACCTGCAATCCCAAGGTTCAATTATGAAACC 578  
 DB 424 AACTCCTTATCTTCAAGCAATTTGAAAACCTGCAATCCCAAGGTTCAATTATGAAACC 483  
 QY 579 ACTGCTCAGAGATATGAGAAAGCTCCGATGGGAAAATTTGATGATTTGTTCTGGATA 638  
 DB 484 ACTGCTCAGAGATATGAGAAAGCTCCGATGGGAAAATTTGATGATTTGTTCTGGATA 543  
 QY 639 GGCACCTGCTGATCAATACAGTGTGCTGAAAATATCTTAAGAGCAGAAATCCGATATA 698  
 DB 544 GGAACAGAGAGCAATAACAGTGTGCTGAAAATATCTTAAGAGCAGAAATCCGATATA 603  
 QY 699 AAGCTGATTTGTTGGAACCACTGTAAGTCCAGTGTCTCAGAGAGAAAGCTGTCCA 758  
 DB 604 AAGCTCTATGTTGGAACCACTGTAAGTCCAGTGTCTCAGAGAGAAAGCTGTCCA 663  
 QY 759 CACAGATTTCAAGGATTTGTTGCTGTTTATCCCTGCTGTCTTGAAGTCAATCTTCTT 818  
 DB 664 CATAGATCCAGGAGATTTGTTGCTGTTTATCCCTGCTGTCTTGAAGTCAATCTTCTT 723  
 QY 819 GATGAGTTGTTCAATATATCAAGTGTGTAAGCAATAGAACTGCAAGCTTCTTGCGCTT 878  
 DB 724 GATGAGTTGTTCAATATATCAAGTGTGTAAGCAATAGAACTGCAAGCTTCTTGCGCTT 783  
 QY 879 AAAGAAGGCTTATTTGTTGGAATATCTTCCGAGCTGCAAGTGTCTGCTTTCAGATT 938  
 DB 784 AAGAAGGCTTATTTGTTGGAATATCTTCCGAGCTGCAAGTGTCTGCTTTCAGATT 843  
 QY 939 GCAAAAAGACCAAAAATGCGGGAGAGCTTATTTGCTGTTTCCCAAGCTTGGGGAG 998  
 DB 844 GCTAAGCGCTTGAATAATGCTGGAGAGCTCATTTGTTTTCCTCAAGCTTGGGGAG 903  
 QY 999 AGGTACCTGCTCCCTGCTGCTATTTGAGTCAAGTGAAGCGGAGCTGAAGCATGACTTT 1058  
 DB 904 CGATATCTTCTGCTGCTGCTGCTTGAATCTTTAGACGGGAGCAGAGACATGACCGTG 963  
 QY 1059 GAGCCCTGAA 1068  
 DB 964 GAGCCTTGA 973

RESULT 4  
 CNAPCCS7  
 LOCUS 1855 bp mRNA linear FLN 28-JAN-2003  
 DEFINITION Citrullus vulgaris mRNA for cysteine synthase, complete cds.  
 ACCESSION D28777.1  
 VERSION D28777.1 GI:466530  
 KEYWORDS O-acetyl-L-serine (thiol)-lyase; O-acetyl-L-serine acetate-lyase (adding hydrogen sulfide); O-acetylserine sulphydrylase.  
 SOURCE Citrullus lanatus (watermelon)  
 ORGANISM Citrullus lanatus  
 Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Cucurbitales; Cucurbitaceae; Citrullus.  
 REFERENCE  
 AUTHORS Noji, M., Murakoshi, I. and Saito, K.  
 TITLE Molecular cloning of a cysteine synthase cDNA from Citrullus vulgaris (watermelon) by genetic complementation in an Escherichia coli Cys- auxotroph  
 JOURNAL Mol. Gen. Genet. 244 (1), 57-66 (1994)  
 MEDLINE 94316193  
 PUBMED 8041362  
 REFERENCE 2 (bases 1 to 1855)  
 AUTHORS Saito, K.

TITLE Direct Submission  
 JOURNAL Submitted (28-FEB-1994) Kazuki Saito, Chiba University, Faculty of Pharmaceutical Sciences, Lab. Mol. Biol. Biotech., Yayoi-cho 1-33, Image-ku, Chiba, Chiba 263, Japan (E-mail: ksaito@p.chiba-u.ac.jp, Tel: 81-43-290-2904, Fax: 81-43-290-2905)  
 COMMENT On Mar 30, 1994 this sequence version replaced gi:391655.  
 D17363: Submitted (05-Aug-1993) to by: Kazuki Saito.  
 FEATURES  
 source  
 1.1855  
 /organism="Citrullus lanatus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3654"  
 /clone="pCCS7"  
 /tissue\_type="etiolated seedlings"  
 564.1541  
 /BC\_number="4.2.99.8"  
 /codon\_start=1  
 /product="cysteine synthase"  
 /protein\_id="BAA05965.1"  
 /db\_xref="GI:540497"  
 /translation="MADAKSTIAKDVTELGNTPLVYLNRVDDCVARVAKLEMEP  
 CSSVKDRIGVSMISDAENKGLITPESVLEPTSGNTGIGAFIAAGRYLLICMPA  
 SMSLEKRTILRAFALVLTDPARGMKGAVQABEIKAKTPNSYLLQEPENPANKH  
 YETTPPEIMRSGSGKIDALVSGIGTGTGAGKYLKEQNPNIKYGVPEVESAILSGA  
 KRPKHIOGIGAGFIPGVLDVNLIDVIVSSESEISITAKLALKEGLLVGISGAA  
 AAAAIRIAKRPENAKLIIVAVPPSPGERIYSTVLFSVKRETNVPEP"  
 ORIGIN  
 Query Match 47.4%; Score 645; DB 8; Length 1855;  
 Best Local Similarity 78.6%; Pred. No. 2.1e-146;  
 Matches 771; Conservative 0; Mismatches 210; Indels 0; Gaps 0;  
 QY 87 ATATGCTGTTGAAAGTCCGGAATTTGCCAAGATGTTACGAATTGATTGTAACC 146  
 DB 561 AAATGCTGATGCAAAAGCTCAACAATAGCTAAAGATGTTACTGAATTTGGAATACA 620  
 QY 147 CCATTACTATATCTAATAAATCTTCCGATGTTGTTGCTCCCGGTTGCTGTAACCTG 206  
 DB 621 CCATTCTATACCTCAACCGTGTGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 680  
 QY 207 GAGTTGATGAGCCATGCTCTAGCTGTAAGGACAGAGATGGGTATAGTATGATGCTGAT 266  
 DB 681 GAGATGATGAGCCCTGCTCCAGTGTCAAGATAGATGCTATAGCATGATTTCAAGAT 740  
 QY 267 GCAGAGAGAGAGGAGCTTATCACACCTGGAAGAGTGTCTCATTTGAGCCAAAGTGGT 326  
 DB 741 GCAGAAAACAAGGCTTTATCACTCCGGGGGAGAGTGTCTTGAATGAACTACCAAGTGGT 800  
 QY 327 AATACTGCAATTGATTTAGCTTCAATGCGAGCAGCCAGGGTTACAAGCTCATATTTACA 386  
 DB 801 AATACTGCTATAGGCTTGGCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 860  
 QY 387 ATGCTGCTTCTATAGTCTTGAAGAGAGATCATTTCTATTAGCTTTTGGAGCTGAGTTG 446  
 DB 861 ATGCTGCTATCAATGAGTCTTGAAGAGAGATCATTTCTATTAGCTTTTGGAGCTGAGTTG 920  
 QY 447 GTTCTGACAGATCTCTGTAAGGAAATGAAGAGTGTGTTTCAAGAGGCTGAAGAGATATTG 506  
 DB 921 GTTCTGACAGATCTCTGTAAGGAAATGAAGAGTGTGTTTCAAGAGGCTGAAGAGATATTG 980  
 QY 507 GCTAAGAGCCCAATGCTTACATCTTCAACAATTTGAAAACCTGCGCAATCCCAAGGTT 566  
 DB 981 GCAAAAACAACCTTAATCTTACATCTTCAACAATTTGAAAACCTGCGCAATCCCAAGGTT 1040  
 QY 567 CATTATGAAACCACTGCTGCAAGATATGAAGAGCTCCGATGGGAAAATTTGATGATTT 626  
 DB 1041 CACTATGAGACCACTGCTGCAAGATCTGAAGAGGCTCAGGTGGGAAAATTTGACGCACTT 1100  
 QY 627 GTTCTGGATAGGCACTGCTGTTACATTAACAAGGCTGGAATAATATCTTAAGAGCAG 686  
 DB 1101 GTCTCTGTTAGGAGCTGAGGTACAGTACAGGCTGCAAGGAAATATCTCAAGAACA 1160

QY 687 AATCCGATATTAAGCTGATTTGCTGGAACCACTTGAAGTCCAGTCTCTGAGAGGA 746  
 Db 1161 AATCCTAATATTAAGTATTAAGCTGGAACCTGTTGAAGTGAATCTATCTGAGAGGA 1220  
 QY 747 AAGCTGTGTCACACAGATTAAGGAGTGTGCTGTTTATCCCTGTGTCTGGA 806  
 Db 1221 AAGCTGTGTCACACATTAAGTGAAGTGTGCTGTTTATCCCTGTGTCTGGA 1280  
 QY 807 GTCAATCTTCTGATGAAGTGTGCTCAATATCAAGTGAAGCAATAGAACTGCAAG 866  
 Db 1281 GTCAATCTTCTGATGAAGTGTGCTCAATATCAAGTGAAGCAATAGAACTGCAAG 1340  
 QY 867 CTCTTGGCGCTTAAGAAGGCTATTGTTGGAATATCTTCCGAGCTGAGCTGCT 926  
 Db 1341 CTGCTTGCATTAAGAAGGCTTACTGTTGATATCATCTGCGCGCGCTGCT 1400  
 QY 927 GCTTTGAGATTTGCAAAAAGACCAAAATGCGGAGCTTATGTTGCGCTTTTCCC 986  
 Db 1401 GCAATTAAGATTTGCAAAAAGACCAAAATGCTGGAAGCTTATGTTGCGCTTTCCC 1460  
 QY 987 AGCTTCGGGAGAGGTAACCTGCTCTCCGCTATTTGAGTCAAGTGAAGCAAGCACTGAA 1046  
 Db 1461 AGTTTGGGAGAGGTAACCTGCTCTCAACCTGCTGTTGAGTCTGTGAAGCAAGCACTGAA 1520  
 QY 1047 AGCATGACTTTTGAGCCCTGA 1067  
 Db 1521 AATATGCTTTTTCAGCCATGA 1541

RESULT 5  
 AB029511 1354 bp mRNA linear PIN 19-JUL-2003  
 LOCUS Solanum tuberosum PCS-1 mRNA for cytosolic cysteine synthase,  
 DEFINITION complete cds.

AB029511  
 VERSION AB029511.2 GI:32997083  
 KEYWORDS  
 SOURCE Solanum tuberosum (potato)

ORGANISM  
 Solanum tuberosum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum.

REFERENCE  
 1 Maruyama, A. and Ishizawa, K.  
 Cytosolic Cysteine Synthase from Potato Sprout  
 Published Only in Database (2001)  
 2 (bases 1 to 1354)  
 Maruyama, A. and Ishizawa, K.  
 Direct Submission  
 Submitted (02-JUL-1999) Akiko Maruyama, Chiba University, Faculty  
 of Pharmaceutical Sciences, Laboratory of Molecular Biology and  
 Biotechnology, Research Center of Medicinal Resources; Yayoi-cho  
 1-33, Inage-ku, Chiba, Chiba 263-8522, Japan  
 (E-mail: amaru@chiba-u.ac.jp, Tel: 81-43-290-2906,  
 Fax: 81-43-290-2905)

REFERENCE  
 AUTHORS Maruyama, A. and Ishizawa, K.  
 TITLE Cytosolic Cysteine Synthase from Potato Sprout  
 JOURNAL Published Only in Database (2001)  
 REFERENCE 2 (bases 1 to 1354)  
 AUTHORS Maruyama, A. and Ishizawa, K.  
 TITLE Direct Submission  
 JOURNAL Submitted (02-JUL-1999) Akiko Maruyama, Chiba University, Faculty  
 of Pharmaceutical Sciences, Laboratory of Molecular Biology and  
 Biotechnology, Research Center of Medicinal Resources; Yayoi-cho  
 1-33, Inage-ku, Chiba, Chiba 263-8522, Japan  
 (E-mail: amaru@chiba-u.ac.jp, Tel: 81-43-290-2906,  
 Fax: 81-43-290-2905)

COMMENT  
 FEATURES  
 source  
 On Jul 19, 2003 this sequence version replaced gi:12081916.  
 Location/Qualifiers  
 1..1354  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /strain="Dansyaku"  
 /db\_xref="taxon:4113"  
 /cell\_type="sprout"  
 /note="tetraploid"  
 1..1354  
 /gene="PCS-1"  
 85..1062  
 /gene="PCS-1"  
 /codon\_start=1  
 /product="cytosolic cysteine synthase"  
 /protein\_id="BAB20861.1"  
 /db\_xref="GI:12081917"  
 /translation="MAGEKIGIAKDVTELIGNTPLVLANVVDGCVARVAARLESMEP"

ORIGIN  
 Query Match 47.2%; Score 643.2; DB 8; Length 1354;  
 Best Local Similarity 76.8%; Pred. No. 6e-146;  
 Matches 786; Conservative 0; Mismatches 238; Indels 0; Gaps 0;  
 CSSVDORIGYMTIDAEKELIPGESVLEPTSGNTGVGLAFMAAKGYKLIITMPS  
 SMSLEIRIILRAFAGELVTDPAKMGKAIISKAEIKAKI PNSTLLOQFENPANPKIH  
 YETTGDEIWKSGNGKVDALVSGTGTGTTTSGSKYLRQNPVTKLXGVEPVBSAILSG  
 GKPGHKIOIGAGFIPGVLEVLIDVVOVSSSESIEMAKLALKEGLLVGSISSGAA  
 AAAAIKVAKR PENAGKLIVIFPSPFERLYLSSVLFETVRREANMTVEP"

QY 47 TCATTCCTGTTCTCAAGTCAAGCTTTGTAGTGAAGCAGATATATATGCTGTGAAGCTC 106  
 Db 42 TCATTCCTGTTCTTGTCAATCCCTAGATTAACCTGATATATATGCGGGGAAAAAGAT 101  
 QY 107 CGGAATTTGCCAAGATGTTAAGGATTTGATTTGTTAAACCCCATTAATATCTAAATPA 166  
 Db 102 TGAATTTGCCAAGATGTTAAGGATTTGATTTGTTAAACCCCATTAATATCTAAATPA 161  
 QY 167 ACTTCGAGATGTTGTTGTTCCCGGTTGCTGCTAACTGAGTTGATGAGCCATGCTC 226  
 Db 162 TGTGTGATGTTGTTGTTGTTCCCGGTTGCTGCTAACTGAGTTGATGAGCCATGCTC 221  
 QY 227 TAGTGTGAAGCAGAGATTTGGTATAGTATGATGCTGATGAGAGAGAGAGGACTTAT 286  
 Db 222 TAGTGTGAAGCAGAGATTTGGTATAGTATGATGCTGATGAGAGAGAGAGGACTTAT 281  
 QY 287 CACACCTGGAAGAGTGTCTCATTTAGGCCAACCAAGTGTATATCTGCAATTGATTAGC 346  
 Db 282 CAAACCCGCGAGAGTGTCTCATTTAGGCCAACCAAGTGTATATCTGCAATTGATTAGC 341  
 QY 347 CTTCATGTCAGCAGCAGCGGCTTACAGCTCAATTAATCAATGCTGCTTCAATGAGTCT 406  
 Db 342 ATTCATGTCGCTGCTTAAAGCTTACCAACTCAATTAATCAATGCTGCTTCAATGAGTCT 401  
 QY 407 TGAGAGAGAGATCAATCTATTTAGCTTTTGAAGCTGAGTTGTTGACAGATCTGCTAA 466  
 Db 402 TGAGAGAGAGATCAATCTATTTAGCTTTTGAAGCTGAGTTGTTGACAGATCTGCTAA 461  
 QY 467 GCGAATGAAGCTGCTGTTCAAGAGGCTGAAGATATTTGGCTAAGAGCCCAATGCTCTA 526  
 Db 462 AGGATGAAGCTGCTGTTCAAGAGGCTGAAGATATTTGGCTAAGAGCCCAATGCTCTA 521  
 QY 527 CATACTTCAAGATTTGAAAAACCTGCAATCCCAAGCTTCATTATGAACAACCTGCTC 586  
 Db 522 TATCTTCAAGATTTGAAAAACCTGCTCAACCAAGATCACTATGAGCAACCTGCTC 581  
 QY 587 AGAGATATGAAGCTCCGATGGAAAAATGATGATTTGTTCTGGATAGGCACTGG 646  
 Db 582 TGAGATCTGAAGAGCTCAATGGAAGAGATGCTAGTCTGGAATTTGGAACAGG 641  
 QY 647 TGGTACAATAAGCTGCTGGAATATCTTAAGAGCAGATCCGAATATTAAGCTGAT 706  
 Db 642 AGGACAGATTAAGCTGCTGGAATATCTTAAGAGCAGATCCGAATATTAAGCTGAT 701  
 QY 707 TGGTGTGAACCAAGTTGAAGTCCAGTGTCTCAGAGAGAAAGCTGTCCACCAAGAT 766  
 Db 702 TGGGTGTGAACCAAGTTGAAGTCCAGTGTCTCAGAGAGAAAGCTGTCCACCAAGAT 761  
 QY 767 TCAAGGATTTGGTGTGTTTATCCCTGCTCTTGAAGTCAATCTTCTGATGAAGT 826  
 Db 762 TCAAGGATTTGGTGTGTTTATCCCTGCTCTTGAAGTCAATCTTCTGATGAAGT 821  
 QY 827 TGTTCAAATATCAAGTATGAAGCAATAGAACTGCAAGCTTCTTGGCTTAAAGAGG 886  
 Db 822 AGTTCAAGTTTCAAGTATGAAGCAATAGAACTGCAAGCTTCTTGGCTTAAAGAGG 881  
 QY 887 CTTATTTTGGGAATATCTTCCGAGCTGCACTGCTGCTTTCAATTTGCAATTTGCAAAAAG 946  
 Db 882 ATTGCTAGTGAATTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 941  
 QY 947 ACCAGAAATGCGGGAAGCTTATTTGCTGCTTTTCCAGCTTCCGAGAGAGTACCT 1006

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
Db	942	CCCTGAAAATGCTGGGAAGCTCATTTGTTGTTATTTTCCCAAGCTTCGGAGAGCCATATCT	1001				
Qy	1007	GTCTCCGCTATTGAGTCAGTAGACGCCGAAGCTGAAGCATGACTTTTGAGCCCTG	1066				
Db	1002	TTCTCTGTGACTCTTCGAAACTGTGCAGACGAGGAGAGAACATGACTGTGGAGCCTG	1061				
Qy	1067	AATT 1070					
Db	1062	AACT 1065					
RESULT 6	ATOACLY	1253 bp	mRNA	linear	PLN 15-FEB-2000		
LOCUS	Arabidopsis thaliana	mRNA for O-acetylserine lyase (At.OAS.5-8).					
DEFINITION	X80376						
ACCESSION	X80376.2	GI:6983573					
VERSION	O-acetylserine(thiol) lyase.						
KEYWORDS	Arabidopsis thaliana (chale cress)						
SOURCE	Arabidopsis thaliana						
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.						
REFERENCE	1	Hell, R., Bork, C., Bogdanova, N., Frolov, I. and Hauschild, R. Isolation and characterization of two cDNAs encoding for compartment specific isoforms of O-acetylserine (thiol) lyase from Arabidopsis thaliana					
AUTHORS	FEBS Lett. 351 (2), 257-262 (1994)						
JOURNAL	94364481						
MEDLINE	8082776						
REFERENCE	2	Hell, R. Direct Submission					
AUTHORS	Submitted (18-JUN-1994) R. Hell, Ruhr-Universitaet Bochum, Lehrstuhl fuer Pflanzenphysiologie, Universitaetsstr. 150, 44801 Bochum, FRG						
JOURNAL	Revised by [4]						
REFERENCE	3	(bases 1 to 1234)					
AUTHORS	Hell, R. Direct Submission						
JOURNAL	Submitted (14-DEC-1994) R. Hell, Ruhr-Universitaet Bochum, Lehrstuhl fuer Pflanzenphysiologie, Universitaetsstr. 150, 44801 Bochum, FRG						
REFERENCE	4	(bases 1 to 1253)					
AUTHORS	Hell, R. Direct Submission						
JOURNAL	Submitted (14-FEB-2000) R. Hell, Ruhr-Universitaet Bochum, Lehrstuhl fuer Pflanzenphysiologie, Universitaetsstr. 150, 44801 Bochum, FRG						
COMMENT	On Feb 16, 2000 this sequence version replaced gi:602329.						
FEATURES	location/Qualifiers						
source	1..1253						
	/organism="Arabidopsis thaliana"						
	/mol_type="mRNA"						
	/cultivar="Columbia"						
	/db_xref="taxon:3702"						
	/clone="At.OAS.5-8"						
	/clone_id="lambda ZAP"						
	/dev_stage="leaf"						
	61..1030						
	/EC_number="4.2.99.8"						
	/codon_start=2						
	/product="O-acetylserine (thiol) lyase"						
	/protein_id="CA56593.2"						
	/db_xref="GI:6983574"						
	/db_xref="GOA:P47998"						
	/db_xref="SWISS-PROT:P47998"						
	/translation="MASRIAKDVTLEIGNTPLVYLNVAEGCVGRVAAKLEMEKPCSSVKDRIGPSMISDAERKGLIKPGESVLIPTSGNTGVLAFATAAKGYKLIITWPASMS TERRIILAFGVELVLTDPAKMKGAIAKAEIILATPNGYMLQOFENPANKIHYET TGEPIWKGTGKIDGFGISGTGTTGAGKYLKEQNANVKLYGVEPVESATISGKKGPHKIQIGAGFIPSTLVNVDLIDEVQVSSDESIMARQALKEGLVGISSGAAAAA						

Query Match	47.2%	Score 642.4	DB 8	Length 1253
Best Local Similarity	79.1%	Pred. No. 9.4e-146		
Matches 763	Conservative 0	Mismatches 201	Indels 0	Gaps 0
QY	105	TCCGGAATTGCCAAGATGTTACGGAATTGATGGTAAACCCCATTAGTATATCTAAAT	164	
Db	68	TCGAGAATTGCTAAGATGTGACTGAATTGATGGACAACCTCATTGGTGTATTGAAC	127	
QY	165	AAACTTGCAGATGTTGTGTGCCCGGTTGCTCTAACTGAGTTGATGAGCCATGC	224	
Db	128	AATGTTGCTGAAGATGTGTGTGTCGTGTGTCTCTAAAGTTGATGAGAACCTGTC	187	
QY	225	TCTAGTGTGAAGCAGAGATTGGGTATAGTATGATTGCTGATGCAGAGAGAGGACTT	284	
Db	188	TCTAGGCTCAAGACAGAGATTGGTTTATGATGATTTCTGATGCAGAAAAGAGGCTTT	247	
QY	285	ATCACACCTGGAAGAGTGTCTCATTGAGCCACAAGTGTATATCTGSCATTGATTA	344	
Db	248	ATCAACCCAGAGAGAGTGTCTGATTGAGCCACAAGTGAACACTGAGTTGGCTTA	307	
QY	345	GCCTTCATGCGACGACGCGAGGGTTACAAGCTCATTAATTAATGCTGCTTATGAGT	404	
Db	308	GCATTACGCGACGCTGCCAAGGCTACAAGCTTATTAATTAATGACGCTTCTATGAGT	367	
QY	405	CTTGAGAGAGAAATCATTTCTATTAGCTTTTGAAGCTGAGTTGGTTCTGACAGATCCTGCT	464	
Db	368	ACTGAGAGAGAAATCATTTCTATTAGCTTTTGAAGTTGAGTTGGTTTAACTGACCCAGCT	427	
QY	465	AAAGGAATGAAGGTGCTGTTCAGAAGCTGAAGAGATATTGGCTAAGACGCCCAATGCC	524	
Db	428	AAAGGCATGAAGAGAGCTATCCCAAGGCGAAGAGATTTGGCGAAAACACCCCAATGCT	487	
QY	525	TACATACTTCAACAATTTGAAAACCCCTGCCAATCCCAAGTTTATTATGAACCACTGCT	584	
Db	488	TACATGCTTCAAGCAATTTGAGAACCCCTGCCAATCCCAAGTTTATTATGAACCACTGCGGA	547	
QY	585	CCAGAGATATGGAAGGCTCCGATGGGAAATTGATGATTTGTTCTGGGATAGGCACT	644	
Db	548	CCTGAGATATGGAAGGCACTGGTGCAAAATCGATGCTTTGTTCTGGGATTTGGTACT	607	
QY	645	GCTGTACAAATAACAGTGTCTGGAATAATCTTAAAGCAGAAATCCGAATATAAGCTG	704	
Db	608	GCTGTACCAATTACAGTGTCTGGAAGTATCTTAAAGAACAGAACGCAACGTCAGCTG	667	
QY	705	ATTGTGTGGAACCAAGTTGAAAGTCCAGTGTCTTCAGAGGAAAGCCTGTCCACACAAG	764	
Db	668	TATGAGTGTGAGCCAGTTGAAAGTGTCTATTCTATCCGTGGGAAAGCCAGTGTCTCAAG	727	
QY	765	ATTCAAGGATTTGTGCTGCTTTTATATCCCTGTGTCTTGAAGTCAATCTTCTGATGA	824	
Db	728	ATTCAAGGATAGAGCTGTGTTTATATCCAGTGTATGATGTGATCTTATGAAGAA	787	
QY	825	GTTGTTCAAAATATCAAGTGAAGCAATAGAAAGTGCAGAACTTCTTGGCTTAAAGAA	884	
Db	788	GTTGTTCAAGTTTCAAGTGAATCCATATGACATGGCAAGGCAAGCTTGTCTTAAAGAA	847	
QY	885	GAGCTATTTTGTGGAATATCTTCCGAGCTGACAGCTGTCTGCTTTTCAGATTGCAAAA	944	
Db	848	GCTCTTCTTGTGGAATATCATCCGGTGCAGCAGCTGTCTGACCAATTAACCTTGACAG	907	
QY	945	AGACCAAGAAAATGCCGGAAGCTTATTGTTGCCCTTTTCCAGCTTCCGGAGAGAGTAC	1004	
Db	908	AGCCCAAGAAAACGCTGGGAAGCTATTGTGGCGATATCCGAGTTTCCGGAGAGAGTAT	967	
QY	1005	CTGTCTCCCGTGTCTATTGAGTCAAGTGAAGCGCAAGCTGAAGCATGACTTTGAGCCC	1064	
Db	968	CTATCAACGGTACTTTTGCATCCGACCAAGGAAGAAAGCGGAAGCCATGACCTTGAAGCT	1027	
QY	1065	TGAA 1068		



Db 1028 TGA 1031

RESULT 7  
ATCYS3A  
LOCUS ATCYS3A 1332 bp mRNA linear PLN 23-AUG-1995  
DEFINITION A thaliana mRNA for cysteine synthase.  
ACCESSION X84097  
VERSION X84097.1 GI:804949  
KEYWORDS cys-3A gene; cysteine synthase.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE  
1 Barroso, C., Vega, J.M. and Gotor, C.  
A new member of the cytosolic O-acetylserine(thiol) lyase gene family in Arabidopsis thaliana  
FEBS Lett. 363 (1-2), 1-5 (1995)

JOURNAL MEDLINE 95246850  
PUBMED 7729527  
REFERENCE 2 (bases 1 to 1332)  
AUTHORS Gotor, C.  
TITLE Direct Submission  
JOURNAL Submitted (25-JAN-1995) C. Gotor, Inst. de Bioquímica vegetal y Fisiología, Dept. de Química, Univ. de Sevilla, A.C.553, E0 41080 SEVILLA, SPAIN

FEATURES  
source  
1.1332  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
1.1332  
/gene="cys-3A"  
99.1067  
/EC\_number="4.2.99.8"  
/EC\_number="4.2.99.8"  
/codon\_start=1  
/product="cysteine synthase"  
/protein\_id="CAA58893.1"  
/db\_xref="GI:804950"  
/db\_xref="GOA:P47998"  
/translation="SWISS-PROT: P47998"  
/translation="MASRIADVTBLIGNTPLVYLNVAAGCVGRAAKLEMMPCSS VKDRIGSMISDAEKGLIKPESVLIPTSGNTGVLAFATAAKGYKLIITPASMSTERRILLAFGEVLTDEPAKMGKAIKAEELIAKTENGYMLQGFENAMPKIHET TGEPIWKGTKIDGFTVSGITGTTGAGTKIKONANVLYGVEPVESAILSGKP GPHKIIGIGAGFIPSVLNDLIDVVOVSDESIDMAROLAKKGLVIGSSGABAAA AIKLAORPENAGKLFVAIFPSFGBRYLSTVLPDTRKEAEAMTFEA"  
234.236  
/gene="cys-3A"  
/bound\_moiety="pyridoxal"

misc\_binding  
234.236  
/gene="cys-3A"  
/bound\_moiety="pyridoxal"

ORIGIN  
Query Match 47.0%; Score 640.8; DB 8; Length 1332;  
Best Local Similarity 79.0%; Pred. No. 2.3e-145;  
Matches 762; Conservative 0; Mismatches 202; Indels 0; Gaps 0;

QY 105 TCCGGAATTGCCAAGATGTTACGAATGATTGTAACCCCATAGTATATCTAAT 164  
DB 105 TCGAGAATTGCTAAAGATGTTGATGTAATGTTGGAACACTTCATTGTTATTTGAC 164  
QY 165 AAATGCGGATGTTGTTGCCCGGCTGCTGCTAAACTGAGTTGATGAGCCATGC 224  
DB 165 AATGTTGCTGAAGATGTTGTTGCTGCTGCTGCTAAGCTTGAGATGATGAGCCGTC 224  
QY 225 TCTAGTGAAGAGACAGATTGGTATAGTATGATGCTGATGACAGAGAGGAGCTT 284  
DB 225 TCTAGGCTCAAGACAGATTGGTATAGTATGATGCTGATGACAGAGAGGAGCTT 284  
QY 285 ATCAACTGGAAGAGATGCTCTCATTTGAGCCAAAGAGTGTATATACTGCATTGATTA 344

Db 285 ATCAACCCAGAGAGATGCTCTGATTGAGCCAAAGTGGAAACACTGAGTTGGCTTA 344

QY 345 GCCTTCATGGCAGACCCAGGGTTACAGCTCATTAATTAATGCTGCTTATAGT 404  
DB 345 GCATTACGGCAGCTGCCAAAGGCTACAAAGCTTATTATTAATGCTGCTTATAGT 404  
QY 405 CTGAGAGAGAAATCATCTTATTAGCTTTTGAGCTGAGTTGTTTGAAGATCTGCT 464  
DB 405 ACTGAGAGAGAAATCATCTTATTAGCTTTTGAGCTGAGTTGTTTGAAGATCTGCT 464  
QY 465 AAGGAATGAAGTGTCTGTTTCAAGAGCTGAGAGATATTGGCTAAGACGCCAATGCC 524  
DB 465 AAGGCATGAAGAGCTATTCGCAAGCGGAGAGATTTTGGCAAAACCCCAATGCT 524  
QY 525 TACATCTTCAACAATTGAAACCTGCAATCCCAAGGTTTATATGAAACCACTGCT 584  
DB 525 TACATGCTTCAAGCTTTGAGAGACCTGCAACCTTAAGATCCACTATGAGTACGGA 584  
QY 585 CCAAGATATGGAAGGCTCCGATGGGAATATGATGATTTGTTCTGGATAGGCACT 644  
DB 585 CCGATATATGGAAGGCACTGCTGGCAATATGATGATTTGTTCTGGATAGGCACT 644  
QY 645 GGTGCTACATATACAGTCTGGAATATATCTTAAGAGCAGATCCGATATTAAGCTG 704  
DB 645 GGTGCTACATATACAGTCTGGAATATATCTTAAGAGCAGATCCGATATTAAGCTG 704  
QY 705 ATTGCTGGAACCAAGTTGAAGTCCAGTCTCTCAGAGAGAAAGCTGCTCACAAG 764  
DB 705 TATGAGTGAAGCCAGTTGAAGTCTATTTATCCGCTGGGAAGCCAGTCTCACAAG 764  
QY 765 ATTCAAGGATGTTGCTGCTTTATCCCTGCTCTTGAAGTCAATCTTCTGATGA 824  
DB 765 ATTCAAGGATGTTGCTGCTTTATCCCTGCTCTTGAAGTCAATCTTCTGATGA 824  
QY 825 GTTGTCAATATCAAGTATGATGAAGCAATAGAACTGCAAGCTTCTTGGCTTAAGAA 884  
DB 825 GTTGTCAAGTATCAAGTATGATGAAGCAATAGCAATGCAAGCTTCTTGGCTTAAGAA 884  
QY 885 GGCCTATTTTGGGAATATCTTCCGAGCTGAGCTGCTGCTTTTGAATGCAAAA 944  
DB 885 GGTCTTCTTGGGAATATATCATCCGTCGAGAGCTGCTGCAATTAATTAATGCAAG 944  
QY 945 AGACCAAGAAATGCGGGAAGCTTATTTGCTGCTTTTCCAGCTTCCGGAAGAGTAC 1004  
DB 945 AGCCAGAAACGCTGGGAAGCTTATTTGCTGCTTTTCCAGCTTCCGGAAGAGTAC 1004  
QY 1005 CTGCTCCGCTGCTATTTGAGTCAAGTGAAGCGGAAGCTGAAAGCATGCTTTGAGCCC 1064  
DB 1005 CTATCAAGCTTCTTTGATGCGCAAGAGAAAGAGCGGAAGCATGCTTTGAGGCT 1064  
QY 1065 TGA 1068  
DB 1065 TGA 1068

RESULT 8  
BJY10845  
LOCUS BJY10845 1280 bp mRNA linear PLN 02-FEB-1999  
DEFINITION Brassica juncea mRNA for O-acetylserine(thiol) lyase, clone OAS-TL4.  
ACCESSION Y10845  
VERSION Y10845.1 GI:2243119  
KEYWORDS O-acetylserine(thiol) lyase.  
SOURCE Brassica juncea  
ORGANISM Brassica juncea  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE  
1 Schafer, H.J., Haag-Kerwer, A. and Rausch, T.  
cDNA cloning and expression analysis of genes encoding GSH synthesis in roots of the heavy-metal accumulator Brassica juncea

L.: evidence for Cd-induction of a putative mitochondrial gamma-glutamylcysteine synthetase isoform  
 JOURNAL Plant Mol. Biol. 37 (1), 87-97 (1998)  
 MEDLINE 98281577  
 PUBMED 9620267  
 REFERENCE 2 (bases 1 to 1280)  
 AUTHORS Schaefer, H.J.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-JAN-1997) H.J. Schaefer, Universitaet Heidelberg, Botanisches Institut, Im Neuenheimer Feld 360, 69120 Heidelberg, FRG

FEATURES  
 source location/Qualifiers  
 1. 1280  
 /organism="Brassica juncea"  
 /mol\_type="mRNA"  
 /sub\_species="Vittasso"  
 /db\_xref="taxon:3707"  
 /clone="OAS-T14"  
 /issue\_type="root"  
 /clone\_lib="lambda ZAP Express"  
 /dev\_stage="mature"  
 74. 1042  
 /codon\_start=1  
 /product="O-acetylserine(thiol) lyase"  
 /protein\_id="CAA71798.1"  
 /db\_xref="GI:2243120"  
 /db\_xref="GOA:O23733"  
 /db\_xref="SWISS-PROT:O23733"  
 /translation="MASRIADVTLLIGNTPLVLTNNVAGVGRVAKLLEMBPSS VKDRIGSMISDABQKLIKPGESVLIETPSGNTGVGLAFTAAAKYKLIITMPASMS VERRIILAFGVELVLTDPKAGMKAIAKEILAKTIPNGVNLQGFENPANPIHYET TGPEIKGTGDKIDGFSVSGITGTTTGAAGKYLKEQNPVLYGVEPIESAILSGKEP GPXKIOGIGAGFIPSLVLEVDLIDEVQVSSDESIDMARLLAKGLLVGISGAATAA AIKAKREPENAKLFAVAVPSPFGRYLYSTVLFDAIRKEASMTFQA"

ORIGIN  
 Query Match 46.7%; Score 636.4; DB 8; Length 1280;  
 Best Local Similarity 78.7%; Pred. No. 2,7e-144;  
 Matches 760; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

QY 105 TCCGAATTGCCAAGATGTACGAATGATGTAACCCCAATGATATCTAAAT 164  
 DB 80 TCTCGAATTGCTAAGATGTACCGAATGATGGGAACACTCCATGGTGTATCTGAAAC 139  
 QY 165 AAACCTGCGGATGTTGTGTTGCCCGGTTGCTGCTAAACTGAGTTGAGCCATGC 224  
 DB 140 AATGTGCTGAAGATGTGTGCTGCTGCTGCTAAGCTCGAGATGATGAAACCTTGC 199  
 QY 225 TCTAGTGAAGACAGATGGGTATAGTATGATGCTGATGACAGAGAGAGGACTT 284  
 DB 200 TCCAGTGTCAAGACAGATGGTATTAGCATGATCTCTGATGACAGAGAGAGGCTCG 259  
 QY 285 ATCAACCTGGAAGAGTGTCTCATTTAGAGCCCAAGTGTATATCTGCAATTGATTA 344  
 DB 260 ATCAACCGAGAGAGAGTGTCTGATTTAGCCCAAGAGTGAACACCGAGTGGGTTA 319  
 QY 345 GCCTTCATGGCAGCAGCGGTTACAGCTCATTAATTAATGACCTTCTATGAGT 404  
 DB 320 GCGTTCACGCGCGCTGCTCAAGGGGTACAGCTCATCTATGACGCTCAATGAGT 379  
 QY 405 CTGAGAGAGAGATCTATTCTATTAGCTTTTGAAGCTGAGTGGTCTGACAGATCTGCT 464  
 DB 380 GTGAGAGAGAGATCTATTCTATTAGCTTTTGAAGTGAAGTGGTCTTAACGACCTGCT 439  
 QY 465 AAGGGAATGAAGGCTGCTTCAAGAGCTGAAGAGATATTGGCTAAGCCCAATGCC 524  
 DB 440 AAGGGGATGAAGGCGCTATCGCAAGGCCGAGAGAGATTTTAGCCAAAACCAATGCT 499  
 QY 525 TACATCTTCAACCAATTGAAAAACCTGCGCAATCCCAAGCTTCAATTATGAACCACTGCT 584  
 DB 500 TACATGCTCCAAACAGTTTGAAGAACCTGCGCAACCCCAAGATACATATGAGACTACTGGA 559  
 QY 585 CCAGAGATATGAAGGCTCCGATGGGAAATGATGCAATTTGTTCTGGGATAGGCACT 644

DB 560 CTTGAGATATGGAAGGCACTGATGAAAAATCGATGATTTGTTCTGAGTTGTTACT 619  
 QY 645 GGTGATCAATTAACAGGCTGCTGAAAAATATCTTTAAGACAGAAATCCGATATAAGCTG 704  
 DB 620 GGTGATCAATTAACAGGCTGCTGAAAAATATCTTTAAGACAGAAATCCGATATAAGCTG 679  
 QY 705 ATGAGTGAAGACAGGTTGAAGAGTCCAGTGTCTTCAGAGAGAAAGCTGTCCACAGCAAG 764  
 DB 680 TATGAGTGAAGACAGGTTGAAGAGTCCAGTGTCTTCAGAGAGAAAGCTGTCCACAGCAAG 739  
 QY 765 ATTCAGGAGATGCTGCTGTTTATCCCTGCTGTCTTGAAGTCAATCTTCTGATGAA 824  
 DB 740 ATTCAGGAGATGCTGCTGTTTATCCCTGCTGTCTTGAAGTCAATCTTCTGATGAA 799  
 QY 825 GTTGTCAATATCAAGTGTGAAGCAATGAAGAACTGCAAAAGCTTCTGCGCTTAAGAA 884  
 DB 800 GTTGTCAAGTTCAGAGTGAATGATCCATGACATGCGCAAGGCTTCTGCTTTAAGAA 859  
 QY 885 GGCCTATTGTGGAATATCTCCGAGCTGACGCTGCTGCTGCTTTTCAAGATTGCAAA 944  
 DB 860 GGTCTTCTGTGGAATATCTCCGAGCTGACGCTGCTGCTGCTTTTCAAGATTGCAAA 919  
 QY 945 AGACCAAGAAATGCGGGAAGCTTATGTTGTCGCTTTTCCAGCTTCCGAGAGAGTAC 1004  
 DB 920 AGGCAAGAAATGCTGGAAGCTTTTGTGCGGCTGCTTCCGAGAGAGTAC 979  
 QY 1005 CTGTCTCCGCTGCTTATGATGATGATGAGACGCGGAAGCTGAAGCAATGCTTTGAGCCC 1064  
 DB 980 CTGTGACTGTACTATTGATGATGATGATGAGACGCGGAAGCTGAAGCAATGCTTTGAGCCC 1039  
 QY 1065 TGAATT 1070  
 DB 1040 TGAATT 1045

RESULT 9  
 SP10AL 1281 bp mRNA linear PLN 02-FEB-1999  
 LOCUS Spinach  
 DEFINITION Spinach mRNA for O-acetylserine(thiol)-lyase, complete cds.  
 ACCESSION D10476  
 VERSION D10476.1 GI:218276  
 KEYWORDS O-acetyl-L-serine acetylase; lyase (adding hydrogen sulfide); O-acetylserine(thiol) lyase; cysteine synthase.  
 SOURCE Spinacia oleracea (spinach)  
 ORGANISM Spinacia oleracea  
 Spinnacia oleracea (spinach)  
 Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Caryophyllales; Amaranthaceae; Spinacia.  
 REFERENCES 1 (bases 1 to 1281)  
 Saito, K., Miura, N., Yamazaki, M., Hirano, H. and Murakoshi, I. Molecular cloning and bacterial expression of cDNA encoding a plant cysteine synthase  
 Proc. Natl. Acad. Sci. U.S.A. 89 (17), 8078-8082 (1992)  
 JOURNAL 92390392  
 MEDLINE 1518833  
 PUBMED 2 (bases 1 to 1281)  
 REFERENCE Saito, K.  
 TITLE Direct Submission  
 AUTHORS Submitted (12-FEB-1992) Kazuki Saito, Chiba University, Faculty of Pharmaceutical Sciences, Lab. Mol. Biol. Biotech.; Yayoi-cho 1-33, Inage-ku, Chiba, Chiba 263, Japan (E-mail:ksaito@p.chiba-u.ac.jp, Tel:043-290-2904, Fax:043-290-2905)  
 JOURNAL Submitted (12-FEB-1992) to DDBJ by:  
 COMMENT Faculty of Pharmaceutical Sciences,  
 Chiba University  
 Yayoi-cho 1-33  
 Chiba 260  
 Japan  
 Phone: 0472-51-1111 x2707  
 Fax: 0472-55-1574.  
 FEATURES  
 source location/Qualifiers  
 1. 1281

```

/organism="Spinacia oleracea"
/mol_type="mRNA"
/db_xref="taxon:3562"
52. .1029
CDS

```

```

/BC_number="4.2.99.8"
/codon_start=1
/product="O-acetylserine(thiol) lyase"
/protein_id="BAA01279.1"
/db_xref="GI:218277"
/translation="MVEKAFIAKDVTELIGKTPLYVLTVAAGCVARVAAKLEGMEP
CSYVDRIIGFSMITDAKSGLITPGESVILIEPTSGNTGIGLAFIAAKGKLIITPA
SMLEERTILRAFGAELITLDPAKGMKAQAKAEIRDKTPNSYILQEPENPANKVH
YETTEPEIWKGTGKIDIFVSGIGTGCTITGAAGKLYKEQNPDVKLIGLEVEESAVLSG
KPGPKHIGLIGLGSFIDPGLVNIIDEVYQISSESIEMAKLALKEGILLVGISSGAA
AAALIAKRPENAGKLIIVAVPSPFGERTLSVLFDSEKAESEMTIES"
1281

```

polYA\_site  
ORIGIN

Query Match	46.7%;	Score 636.4;	DB 8;	Length 1281;
Best Local Similarity	77.7%;	Pred. No. 2.7e-144;		
Matches 769; Conservative	0;	Mismatches 221;	Indels 0;	Gaps 0;

QY	87	ATATATGCGCTGTTGAAGGTCGCCGAATTTGCCAAGATGTTAACGAATTGATTTGGTAAACC	146
Db	49	AAATATGTTGAGGAGAAAGGCGCTTCATTGCTAAAGATGTGACTGAATTGATTGGGAAAACG	108
QY	147	CCATTAGTATATCTAAATAAACTTCCGGATGTTGTGTTCGCCGGGTGCTGCTAAACTG	206
Db	109	CCATTGTATATCTCAACACTGTGCGCGATGTTGTGTTCCTGTTGCTGCAAGCTTG	168
QY	207	GAGTTGATGAGCCATGCTCTAGTGTGAAGGACAGATTTGGGTATAGTATGATTGCTGAT	266
Db	169	GAAAGATGGAACCTTGTCTAGTGTTAAGACAGATTTGGGTTCAGTATGATTACTGAT	228
QY	267	GCAGAGAGAAAGGCACTTATCAACCTGGAAGAAGTGTCTCATTTAGCCACAACTGCT	326
Db	229	GCTGAAAAAAGCGGGCTTATTAACCTGAGAGAGTGTCTGATTGAGCCACACAGTGA	288
QY	327	AATACTGGCATTGATTAAGCTTCACTGAGCAGACCCAGGGTTACAAGCTCATTAATTACA	386
Db	289	AATACTGGCATTGATTAAGCTTCACTGAGCAGACCTAAAGTTACAAGCTCATTAACG	348
QY	387	ATGCGTGTCTTATAGAGCTTGAGAGAGAAATCATTTCTTAATAGCTTTTGAGCTGAGTTG	446
Db	349	ATGCCAGCATCAATGAGTTTGAGAGCGAGACTATTCTCAGGGCTTTGTTGCTGAGCTT	408
QY	447	GTTCTGACAGATCTGCTAAGGGAATGAAGGTGCTGTTCAGAAAGCTGAAGAGATATG	506
Db	409	ATCCTTACTGATCCAGCAAAAGTATGAAGGGGCTGTTCAGAAAGCTGAGAGATCCGT	468
QY	507	GCTAAGACGCCCAATGCTCTACATCTTCAACAATTTGAAAAACCTGCCATCCCAAGTT	566
Db	469	GACAAAACCTCCTAATTCTATATATACTACAACAGTTTGAAAAACCTGCCAACCCAAAGTT	528
QY	567	CATTATGAACAACCTGTGTCCAGAGATATGGAAGGCTCCGATGGGAAAATTGATGCATT	626
Db	529	CATTATGAACAACCTGTGTCCAGAAATTTGGAAGGCAACAGGTGAAAAAATTGATATATTC	588
QY	627	GTTTCTGGGATAGGCACGTGTGTGTAACAATACAGGTGCTGGAAAATATCTTAAAGAGCAG	686
Db	589	GTTCTGTGAATAGGAGACTGAGAGGTACAATACAGGTGACGAAAAATACCTAAGGAAACA	648
QY	687	AATCCGATATAAAGCTGATTGTGTGGAACCAAGTTGAAGTCCAGTGTCTCCAGAGGA	746
Db	649	AAACCGGATGTTAAGCTTAATTGGCTTGAAACCAAGTGAAGAGTCTGTATTTGTCTGAGGA	708
QY	747	AAGCTGTGTCACACAAGATTCAAGGGATTGTGTGTTTTATCCCTGTGTCTTGGA	806
Db	709	AAACCTGGCCACATTAAGATTCAAGGACTTGAGCTGGAATTCATACCTGTGTCTTGAT	768
QY	807	GTCATCTTCTTGATGAAGTGTTCAAATATCAAGTGAAGCAATAGAACTGCAAG	866
Db	769	GTCATATATTATGATGAAGTGTTCAGATATCCAGTGAAGAACTTAATTGAATGCGCCAA	828

Oy	867	CTTCTTGCGCTTAAGAAGGCCATTGTGTCGAATACTTTCCGGAGCTGCAGCTGCTGT	926
Dd	829	TTGCTCGCCCCCAAGGAAGGCTCTACTGTGTGGAAATTCATCTGTGTCTGTCTGTCCGCT	888
Oy	927	GCTTTTCAGATTGCCAAAAAGACCAGAAAATGCCCCGGAAGCCTTATTGTTGCCGTTTTCCC	986
Dd	889	GCCATTAAAGTGGCAAAGAGGCCCTGAAAATGCTGGAAAACTCATGTGCTGTCTTTCCC	948
Oy	987	AGCTTCGGGAGAGGTACCTGTCTCCGTCGTATTGAGTCAGTAGAAGCGAAGCTGAA	1046
Dd	949	AGCTTTGGCGAACGATATTATCTCTCGGTGTGTTTGATTCAGTGAGGAAGAGCGAGAG	1008
Oy	1047	AGCATGACTTTTGAGCCCTGAATTCGCCGT	1076
Dd	1009	AGCATGTTATTGAGTCTCTAAGTTCCAGTT	1038

	RESULT 10				
LOCUS	E06716				
DEFINITION	CDNA encoding Spinacia cysteine synthetase.	1303 bp	RNA	linear	PAT 29-SEP-1997
ACCESSION	E06716				
VERSION	E06716.1	GI:2174898			
KEYWORDS	JP 1994038770-A/1.				
SOURCE	Spinacia oleracea (spinach)				

ORGANISM Spinacia oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Caryophyllales; Amaranthaceae; Spinacia.  
 1 (bases 1 to 1303)  
 Saitou, K. and Murakoshi, I.  
 GENE CODING CYSTEINE SYNTHETASE  
 Patent: JP 1994038770-A 1 15-FEB-1994;

**COMMENT**

```

PN      JP 1994038770-A/1
PD      15-FEB-1994
PF      05-FEB-1992 JP 1992020315
PI      SAITOU KAZUSUE, MURAKOSHI ISAMU
PC      C12N15/60, C12N9/88, (C12N9/88, C12R1:19);
CC      strandedness: Double;
CC      topology: Linear;
CC      hypothetical: No;
CC      anti-sense: No;
CC      *source: strain=parade;
CC      *source: clone=PTV118N;
FH      Key
FH      Location/Qualifiers
FT      5'UTR
FT      CDS
FT      1..51
FT      52..1029
FT      /product='Spinacia cysteine synthetase' FT
3'UTR   1030..1030.

```

```

FEATURES
    source
        location/Qualifiers
            1..1303
                /organism="Spinacia oleracea"
                /mol_type="Genomic RNA"
                /db_xref="taxon:3562"
ORIGIN

```

Query Match	46.7%;	Score 636.4;	DB 6;	length 1303;
Best Local Similarity	77.7%;	Pred. No. 2.7e-144;		
Matches 769;	Conservative 0;	Mismatches 221;	Indels 0;	Gaps 0;
QY	87	ATAATGCGCTGTGAAGCTCCGGAATGCCAAGAATGTTACGGAATTGATTGTTAAACC	146	
Db	49	AAATGTTGAGGAGAGAGCCCTTCATTGCTAAAGATGATGACTGAATTGATTGGGAAACG	108	
QY	147	CCATTAGTATATCTAATAATAACTGCGGATGTTGTGTGCCCCGCGTGTGCTCTAAACTG	206	
Db	109	CCATTGTTATATCTCAACACTGTGCGCGAATGTTGTGTGCTCGTGTGCTGCAAAAGCTG	168	
QY	207	GAGTTGATGGAGCCATGCTCTAGTGTGAAGACAGATTGGGTATAGTATGATTGCTGAT	266	



Db 169 GAAGGAATGAACTTGTCTAGTGTAAAGACAGATTGGGTTGAGTATGATTACTGAT 228  
QY 267 GCAGAGAGAGAGAGCTTATCACAACCTGGAAAGAGTCTCTCATGAGCCCAACAGTGT 326  
Db 229 GCTGAAAAAAGCGGCTTATTACACCTGGAGAGTGTCTGATTGAGCCCAACAGTGA 288  
QY 327 AATACGTCATTTGATTAGCTTCAATGACAGAGCCAGGGTTACAAGCTCATTAATTACA 386  
Db 289 AATACGTCATTTGATTAGCTTCAATGACAGAGCCAGGGTTACAAGCTCATTAATTACG 348  
QY 387 ATGCTGCTTCTATGAGTCTTGAAGAGAGATCATTTATTAAGCTTTGAGCTGAGTTG 446  
Db 349 ATGCCAGCATCAATGAGTCTTGAAGGAGAGACTATTCTCAGGCTTTGCTGAGCTT 408  
QY 447 GTTCTGACAGATCTCTGCTTAAAGGAATGAAAGTGTCTTCAAGAGGCTGAAGATATTG 506  
Db 409 ATCCTTACTGATCCAGCAAAAGGATGAAAGGCTGTTCAGAGGCTGAGAGATCCGT 468  
QY 507 GCTAAGACGCCCAATGCTTACATCTTCAACATTTGAAAACCTGCCAATCCCAAGTT 566  
Db 469 GACAAACTCTTAATTCATATATATCTACCAACAGTTGAAAACCTGCCAATCCCAAGTT 528  
QY 567 CATTATGAAACCACTGCTCCAGAGATATGGAAGGCTCCGATGGGAAATTTGATGATTT 626  
Db 529 CATTATGAAACCACTGCTCCAGAGATATGGAAGGCTCCGATGGGAAATTTGATGATTT 588  
QY 627 GTTCTGAGATAGGCACTGCTGTGTAATTAACAGTGTCTGGAATATCTTAAAGAGCAG 686  
Db 589 GTCTCTGAAATAGGCACTGCTGTGTAATTAACAGTGTCTGGAATATCTTAAAGAGCAG 648  
QY 687 AATCCGATATTAAGCTGATTTGCTGGAACCACTGGAAGTCCAGTCTCTCAGAGAGA 746  
Db 649 AATCCGATATTAAGCTGATTTGCTGGAACCACTGGAAGTCCAGTCTCTCAGAGAGA 708  
QY 747 AAGCTGTGTCACACAGATTCAGAGGATTTGCTGCTTTTATCCCTGCTGCTTGA 806  
Db 709 AAGCTGTGTCACACAGATTCAGAGGATTTGCTGCTTTTATCCCTGCTGCTTGA 768  
QY 807 GTCAATCTTCTGATGAGTGTTCATTAATATCAAGTATGAGCAATAGAACTGCAAG 866  
Db 769 GTCAATCTTCTGATGAGTGTTCATTAATATCAAGTATGAGCAATAGAACTGCAAG 828  
QY 867 CTTCCTGCTTAAAGAGGCTTATTTGCTGGAATATCTTCCGAGCTGCAAGCTGCT 926  
Db 829 TTGCTGCTTCAAGAGGCTTATTTGCTGGAATATCTTCCGAGCTGCAAGCTGCT 888  
QY 927 GCTTTTCAAGTTCAGAAAGAGCAAGAAATGCGGAGAGCTTATTTGCTGCTTTTCC 986  
Db 889 GCTTTTCAAGTTCAGAAAGAGCAAGAAATGCGGAGAGCTTATTTGCTGCTTTTCC 948  
QY 987 AGCTTCGGGAGAGTACCTGCTCTCCGCTGCTATTTAGTCAAGAGAGAGAGCTGAA 1046  
Db 949 AGCTTCGGGAGAGTACCTGCTCTCCGCTGCTATTTAGTCAAGAGAGAGAGAGAG 1008  
QY 1047 AGCATGACTTTTGAAGCCCTGAATTCCTGTT 1076  
Db 1009 AGCATGACTTTTGAAGCCCTGAATTCCTGTT 1038

## RESULT 11

AF044172

1308 bp

mRNA

linear

PLN 06-JUL-1998

## LOCUS

AF044172

Solanum tuberosum cysteine synthase mRNA, complete cds.

## DEFINITION

AF044172

Solanum tuberosum cysteine synthase mRNA, complete cds.

## ACCESSION

AF044172.1

GI:3290019

## VERSION

AF044172.1

GI:3290019

## KEYWORDS

Solanum tuberosum (potato)

## SOURCE

Solanum tuberosum

## ORGANISM

Solanum tuberosum

## REFERENCE

1 (bases 1 to 1308)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Solanales; Solanaceae; Solanum.

AUTHORS Hesse, H. and Hoefgen, R.  
TITLE Isolation of cDNAs encoding cytosolic (Accession No. AF044172) and  
plasticidic (Accession No. AF044173) cysteine synthase isoforms from  
Solanum tuberosum (PGR98-057)  
JOURNAL Plant Physiol. 116, 1604 (1998)  
REFERENCE 2 (bases 1 to 1308)  
AUTHORS Hesse, H.  
TITLE Direct Submission  
JOURNAL Submitted (22-JAN-1998) Freie Universitaet Berlin, Institut fuer  
Angewandte Genetik, Albrecht-Thaer-Weg 6, Berlin 14195, Germany  
FEATURES  
source location/Qualifiers  
1..1308  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultiivar="Berolina"  
/db\_xref="taxon:4113"  
35..1012  
/note="CS-A; O-acetylserine (thiol) lyase; cytosolic  
isoform"  
/codon\_start=1  
/product="cysteine synthase"  
/protein\_id="AAC25635.1"  
/db\_xref="GI:3290020"

## CDS

/translation="MAGEKIGIAKDVTELLIGTEPLVYLNVDGCVARVAALIESMEP  
CSVKDRIGVSMITDAERKGLIKPESVLIETSGNTGVGLAFMAAAGKYLITPES  
SMLEKRIILRGFSELVITDPKMGKAIKAEIKAKTENSYLIOQFENPAPKIH  
YETGPRIWKSNGKYDALASGIGTGITGSGKYLREQNPNVKLYGEPVESAIISSGAA  
KRPPIKIGIAGFIPGVLEVNLIIDVVOVSSESIEMAKLALKEGLVGISSGAA  
AAAIKVAKRPENAKLIVIVIPSPGERYLSVLFETVREAENMTVEP"

## ORIGIN

Query Match 46.5%; Score 632.8; DB 8; Length 1308;  
Best Local Similarity 76.9%; Pred. No. 2e-143;  
Matches 772; Conservative 0; Mismatches 232; Indels 0; Gaps 0;  
QY 67 AAGCTTTGATGAGCAGATATATGCTGTGAAAGGTCGGAATTGCCAAGATGTTA 126  
Db 12 AATCCTGAAATTAACCTGATATATGCGGGGAAAAAGATTGGAATTGCCAAGATGTTA 71  
QY 127 CGGAATTGATGTTAAACCCTTAGATATCTAAATAAATCTGCGAGTGTGTG 186  
Db 72 CTGAATTGATGTTAAACCCTTAGATATCTAAATAAATCTGCGAGTGTGTG 131  
QY 187 CCCGGTCTGCTAACTGAGTGTATGAGAGCCATGCTTGTGTAAGAGCAGATG 246  
Db 132 CACGTGTGCTGCCAAGCTGGAAGAGAGAGCCATGCTTGTGTAAGAGCAGATG 191  
QY 247 GGTATGATGATTTGCTGATGCAAGAGAGAGAGGACTTATCACTGGAAGAGTGC 306  
Db 192 GTTATGATGATTTGCTGATGCAAGAGAGAGAGGACTTATCACTGGAAGAGTGC 251  
QY 307 TCAATGAGCCAAACAAGTGTATATGAGGATTTGATTAAGCTTATGAGAGAGG 366  
Db 252 TCATGAGAGCCAAACAAGTGTATATGAGGATTTGATTAAGCTTATGAGAGAGG 311  
QY 367 GTTACAGCTCATTAATTACATGCTCTTCTATGAGTCTTGAAGAGAGATCTAT 426  
Db 312 GCTACAACTCATTAATTACATGCTCTTCTATGAGTCTTGAAGAGAGATCTAT 371  
QY 427 TAGCTTTGAGAGCTGAGTGTCTGACAGATCTCTGTAAGGAGATGAAGGTCTTC 486  
Db 372 GTGCTTTGAGAGCTGAGTGTCTGACAGATCTCTGTAAGGAGATGAAGGTCTTC 431  
QY 487 AGAAGCTGAAGAGATTTGGCTAAGAGCCCAATGCTCTACATCTTCAACATTTGAAA 546  
Db 432 CAAAGCTGAAGAGATTTGGCTAAGAGCCCAATGCTCTACATCTTCAACATTTGAAA 491  
QY 547 ACCCTGCCAATCCCAAGGTTCAATATGAAAACCACTGCTCCAGAGATATGAAAAGGCTCCG 606  
Db 492 ACCCTGCCAATCCCAAGGTTCAATATGAAAACCACTGCTCCAGAGATATGAAAAGGCTCAA 551  
QY 607 ATGGAAATTTGATGATTTGTTCTGAGATGAGCACTGCTGTGATCAATTAAGGTCTG 666

Db 552 ACCGGAAGTGATGCTCTAGCCTCTGGAATTGGAACAGAGGACGATTAAGTCTGAC 611

Qy 667 GAAATATCTTAAAGACGAATCCGAATATAAGTGAATGCTGTGGAACCGTTGAAA 726

Db 612 GCAAGTATTGAGAGAGCAACCCCAACGTTAGCTGTATGCGGTGGAACCAAGTTGAAA 671

Qy 727 GTCCAGTCTCTCAGAGAGAAAGCCTGTGTCACACAGATTCAGGGGATGCTGCTGT 786

Db 672 GTGCTATCCTTTCTGTGGAAGCCTGTGTCACATAGATTCAGGGGATGCTGCTGT 731

Qy 787 TTATCCCTGCTGTCTTGAAGTCAATCTTGTGATGAGTGTTCATATATCAAGTATG 846

Db 732 TCATTCCTGCTGTGTTGGAAGTAACTTATGATGATGTAGTTCAGTTTCAAGTGAAG 791

Qy 847 AAGCAATAGAACTGCAAGCTTCTTGCCTTAAAGAGCCTATTGTGGAATATCTT 906

Db 792 AATCCATAGAAATGGCTAGCTTCTGCAATGAAGAGGATGCTAGTGGGAATTCAT 851

Qy 907 CCGAGCTGACAGCTGCTGCTCTTTCAGATTCGAAAGAACAGCAAGAAATGCCGGAAGC 966

Db 852 CTGCTGCTGCTGCGCGCGGCAATTAAGTGTCTAAGCGCCCTGAAAATGCTGGGAAGC 911

Qy 967 TTATGTGTCGCTTTTCCCACTTGGGGAGAGGTAACCTGCTCCCTGCTATTGAGT 1026

Db 912 TCATTTGTTTATTTTCCCAAGCTTCGAGAGCGATATCTTCTCTGATCTCTCGAAA 971

Qy 1027 CAGTGAAGCGGAAGCTGAAACCATGACTTTTGAGCCCTGAAT 1070

Db 972 CTGTCAAGAGAGAGAGCAACATGACTGTGAGCCTTGAAT 1015

RESULT 12

LOCUS AY063827 1233 bp mRNA linear PLN 23-SEP-2002

DEFINITION Arabidopsis thaliana At3g22460 mRNA sequence.

ACCESSION AY063827

VERSION AY063827.1 GI:17380705

KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 1233)

Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shin, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Arabidopsis Full Length cDNA Clones

TITLE Arabidopsis Full Length cDNA Clones

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1233)

Yamada, K., Ban, J., Banno, F., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Meyer, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shin, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Direct Submission

TITLE Submitted (19-NOV-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

JOURNAL 3 (bases 1 to 1233)

REFERENCE Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shin, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

TITLE Direct Submission

JOURNAL Submitted (23-SEP-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shin, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

FEATURES

source

1. 1233

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="3"

/clone="RAFL06-87-B19 (R11881)"

/ecotype="Columbia"

/note="This clone is in a modified pBluescript vector (Lambda ZAP) as a XhoI/SacI insert."

gene

1. 1233

/gene="At3g22460"

misc\_feature

84. 1061

/gene="At3g22460"

/note="This may be a potential product of a partially processed transcript corresponding to gene At3g22460."

/evidence=experimental

misc\_difference

1216

/gene="At3g22460"

/note="not present in genomic sequence"

ORIGIN

Query Match 46.4%; Score 632.6; DB 8; Length 1233;

Best Local Similarity 77.1%; Pred. No. 2.3e-143;

Matches 770; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

Qy 71 TTGTAGTGAAGCAGATATATATGCTGTGGAAGGTCGCCGAATGCCAAGATGTACGGA 130

Db 65 TTTTCAAGTGAAGCTTGAACATGCTTCAGTGGCTCTTAATCGCTTAAGATGTACTGA 124

Qy 131 ATTGATGTAAGAACCCCATTAATATCTTAATTAACCTGCGGATGCTGTGCTCCG 190

Db 125 ACTAATAGGAACACTCCATTAGTATCTTAACAAAGTAGCTTAAGATTGTGTGCTCA 184

Qy 191 GGTGCTGTAAGCTGAGTGTGATGAGGACCTCTAGTGTGAAGACAGAGATTGGGTA 250

Db 185 TGTGCTGTAAGCTTGAAGTATGAGAACCATGTTCTAGCGCTAAGACAGAGATTGGTGA 244

Qy 251 TAGTATGATGCTGATGAGAGAGAGGAGGACTTATCACACCTGGAAGAGTGTCTCAT 310

Db 245 TAGTATGATGCTGATGAGAGAGGAGGACTTATTAACCGGAGAGAGTGTGTGAT 304

Qy 311 TGAGCCAAAGTGTATATAGCTGATGAGTATGAGCTTCAAGGAGAGAGGAGGTTA 370

Db 305 TGAGCCAAAGTGTGAGAGAGGAGGAGTGTGATGATTAAGCTGAGAGAGGAGGCTA 364

Qy 371 CAAGCTCAATTAATCAATGCTGCTTATGAGTCTTGAAGAGAGATCATTTATTAGC 430

Db 365 TAAGCTGTTATTAATCAATGCTGCTTATGAGTCTTGAAGAGAGATCATTTATTAGC 424

Qy 431 TTTTGAAGTGAAGTGTGCTGAGAGATCTGCTAAGGAGAGAGGAGTGTGCTGAGAA 490

Db 425 TTTTGAAGTGAAGTGTGCTGAGAGATCTGCTAAGGAGAGAGGAGTGTGCTGAGAA 484

QY 491 GGCTGAAGATATTGGCTAAGACGCCCAATGCTTACATCTTCAACAATTGAAACC 550  
DB 485 GGGGAGAGATTTTGGCGAAGAACCTTAATGTTATATGCTTCAACGTTTGAAGATCC 544  
QY 551 TGGCAATCCCAAGGTTTATTATGAAGCCACTGGTCCAGAGATATGAAAGGCTCCGATGG 610  
DB 545 TGGTATCCCAAGATCCATTATGAGACTAGTGAAGCTTGAATATGAAAGGTTCTGGTGG 604  
QY 611 GAAATTTGATGATTTGTTTCTGGGATAGCACTGGTGTACAAATACAGGCTGTGAAA 670  
DB 605 AAAAGTGATGCTTTGTTTCTGGTATGTTAGTGTGTGTTACCTGAACAGGCTGTGAAA 664  
QY 671 ATATCTTAAGAGCAGAAATCCGAATATAAAGCTGATTTGGTGAAGCAGTGAAGTCC 730  
DB 665 GTATCTCAAGAGACAGAACACAAACATAAGCTGATGTGTGAGCTGTGTGAAGGCC 724  
QY 731 AGTGTCTCAGAGAGAACCTGTGTCACACAGATTCAGGAGTGTGCTGTTTAT 790  
DB 725 TATTTCTGTCGGTGAAGCAGGTCCTCCACAAATTCAGGTATAGTGTGCTGTTTAT 784  
QY 791 CCTGTGTCTTGAAGTCAATCTTCTTGAATGATTTGTCAAATATCAAGTATGAAGC 850  
DB 785 TCCAGGATTTTGAATGTGATCTTATAGATGAAGTGTTCAGGTTCAAGTGAAGATC 844  
QY 851 AATAGAACTGCAAGCTTCTTGGCTTAAGAGAGCTTATTTGGGATATCTTCCGG 910  
DB 845 CATGACATGGCAGGCTTCTGCTAGGAAAGAGGCTGTGTAGTGGAAATCTCATCTGG 904  
QY 911 AGCTGACGCTGCTGCTCTTTCAGATTCGAAAAGAACAGAAATGCGGGAAGCTTAT 970  
DB 905 AGCAGCAGCTACTGCAACATCAACTGCAAGAGAGGCCAGAAATGCTGGGAAGCTTAT 964  
QY 971 TGTGCGCTTTTCCAGCTTCGGGAGAGTACTGCTCCGCTATTTGAGTCAGT 1030  
DB 965 AGTGGCGGTGTTCCGAGTTTCGGGAAAGATATCTATGACTGTACTGTTGATGACGC 1024  
QY 1031 GAGACGCGAAGCTGAACATGACTTTGAGCCCTGAAT 1069  
DB 1025 AAGGAAGAGCGCAGACTATGACTTTGAGCCCTGAAT 1063

RESULT 13  
LOCUS AY045825 1254 bp mRNA linear PLN 18-SEP-2002  
DEFINITION Arabidopsis thaliana putative cytosolic O-acetylserine(thiol)lyase  
(At4g14880) mRNA, complete cds.  
ACCESSION AY045825 GI:15027936  
VERSION AY045825.1  
KEYWORDS FLI\_CDNA.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1 (bases 1 to 1254)  
Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K.,  
Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,  
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,  
Satou, M., Seki, M., Shim, P., Southwick, A., Shinzaki, K.,  
Davis, R.W., Ecker, J.R. and Theologis, A.  
Arabidopsis Full length cDNA clones  
Unpublished  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
2 (bases 1 to 1254)  
Yamada, K., Liu, S.X., Pham, P.K., Banh, J., Dale, J.M., Goldsmith, A.D.,  
Jiang, P.X., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,  
Toriumi, M., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P.,  
Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A.,  
Karlin-Neumann, G., Kawai, J., Kim, C., Koesema, E., Lam, B., Lin, J.,  
Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,  
Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Tracy, S.E.,

TITLE  
JOURNAL  
COMMENT

Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
Direct Submission  
Submitted (10-JUL-2001) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN  
Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,  
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the  
sequencing and annotation of the RAFL cDNAs: Yamada, K., Liu, S.X.,  
Pham, P.K., Banh, J., Dale, J.M., Goldsmith, A.D., Jiang, P.X.,  
Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M.,  
Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T.,  
Karlin-Neumann, G., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C.,  
Miranda, M., Nguyen, M., Palm, C.J., Shim, P., Southwick, A.,  
Tracy, S.E., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to  
this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC)  
contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis  
genome submitted to GenBank.

FEATURES  
source

location/Qualifiers

1..1254

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="4"

/clone="RAFL05-20-P15 (R10445)"

/ecotype="Columbia"

/note="This clone is in a modified pBluescript vector  
(lambda ZAP) as a XhoI/SstI insert."

1..1254

/gene="At4g14880"

1..77

/gene="At4g14880"

78..1043

/gene="At4g14880"

/note="contains cysteine synthase/cystathionine  
beta-synthase P-phosphate attachment site AA35-53"

/codon\_start=1

/evidence="experimental"

/product="putative cytosolic O-acetylserine(thiol)lyase"

/protein\_id="AAK76499.1"

/db\_xref="GI:15027937"

/translation="MASRIADVTTELIGNTPLVNNVABGCVGRVAALKEEMPCSSV  
KDRIGFSMSDAERKGLIKPGBSVLIBPTSGNTGVGLAFTAAAGYKLIITPASMST  
ERRITLAFGEVLELTPAKGKGAIAKABEILAKTPNGYMLQOFENPANKIHYET  
GPEIWKGTGKIDFVSGIGTGTITGAGKYLEGNANVKLYGVEPVESAIIISGKPG  
PKIIGIAGFIPSVLNVLDIDEVQVSSDESIDMARQALKEGLVGISSGAAAAA  
IKLAQRPENAGLEVALFPFSGERYLSTVLPDATRKAEMTFEA"

134..135

/gene="At4g14880"

/note="compared to genomic sequence"

/replace="gtrac"

136..137

/gene="At4g14880"

/note="compared to genomic sequence"

/replace="ttg"

1044..1254

/gene="At4g14880"

1237

/gene="At4g14880"

/note="not present in genomic sequence"

misc\_difference

misc\_difference

3'UTR

misc\_difference

misc\_difference

misc\_difference

misc\_difference

misc\_difference

misc\_difference

misc\_difference

misc\_difference

misc\_difference

misc\_difference

misc\_difference

misc\_difference

misc\_difference

misc\_difference

misc\_difference

misc\_difference

misc\_difference

misc\_difference

misc\_difference

## ORIGIN

Query Match 46.1% Score 628; DB 8; Length 1254;  
Best Local Similarity 78.9%; Pred. No. 3e-142;  
Matches 761; Conservative 0; Mismatches 200; Indels 3; Gaps 1;







QY 969 ATGTGGCGTTTTCCTCCAGCTTGGGGAGAGGTACTGTCTCTCCGTGCTATTGAGTCA 1028  
Db 781 ATGTGGCGATCTTCCCTAGTTTGAGAGCGGTACTTAATCTCTGTGCTTTTGAATCA 840  
QY 1029 GTGAGACGCGAAGCTGAAGCATGACTTTGAGCCCTGAATTCCGTTTAAGGCTCTCA 1087  
Db 841 GTGAAGAAGAGGCGCGAGACATGTTTGTAGCCTTGAGGTGAAGAACAATGCTTTCA 899

Search completed: May 21, 2004, 20:33:04  
Job time : 5588 secs







PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.

PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149175P.  
PR 17-AUG-1999; 99US-0149368P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 23-AUG-1999; 99US-0150566P.  
PR 25-AUG-1999; 99US-0150884P.  
PR 26-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 27-AUG-1999; 99US-0151303P.  
PR 30-AUG-1999; 99US-0151438P.  
PR 31-AUG-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.



PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 47.1%; Score 641.4; DB 3; Length 1120;  
 Best Local Similarity 79.1%; Pred. No. 6.8e-169;  
 Matches 762; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

QY 105 TCCGGAATGCCAAGATGTTACGGAATGTTGTTAAACCCCAATTAATCTAAT 164  
 DB 158 TCGAATGCTAAAGATGTAATGTAATGGAACACTCCATTTGTTATTTGAAC 217  
 QY 165 AAACCTGCGAGTGTGTGTTGCTCCGCTGCTGCTAACTGAGTGAAGCCATGC 224  
 DB 218 AATGTTGCTGAAGATGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 277  
 QY 225 TCTAGTGAAGACAGATGTTGTTAGTATGATGTTGTTGTTGTTGTTGTTGTT 284  
 DB 278 TCTAGCTGAAGACAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 337  
 QY 285 ATCACACCTGGAAGAGTGTCTCTATGAGCCCAAGTGTGTTAATCTGCAATGATTA 344  
 DB 338 ATCAACCAAGAGAGTGTCTCTATGAGCCCAAGTGTGTTAATCTGCAATGATTA 397  
 QY 345 GCCTTCATGAGCAGCCAGGCTTACAGCTCATTAATTAATGCTGCTTCTATGAGT 404  
 DB 398 GCATTCACGCACTGCCAAGGCTTACAGCTTATTAATTAATGCACTTCTATGAGT 457  
 QY 405 CTTGAGAGAAATCATTTCTATTAGCTTTGAGCTGAGTGTGTTGTTGTTGTTGTT 464  
 DB 458 ACTGAGAGAAATCATTTCTATTAGCTTTGAGCTGAGTGTGTTGTTGTTGTTGTT 517  
 QY 465 AAGGAATGAAGAGTGTCTCTGAGAGGCTGAAGATGTTGCTAAGACGCCCAATGCC 524  
 DB 518 AAGGCATGAAGAGCTATCGCAAGGCGAAGATTTGGCGAAAAACCAATGCT 577  
 QY 525 TACATACCTCAATTTGAAAAACCTGCCAATCCCAAGGTTGTTATGAAACCACTGCT 584  
 DB 578 TACATGCTTCAAGCTTTGAGAACCTGCCAATCCCAAGGTTGTTATGAAACCACTG 637  
 QY 585 CCAGAGATGAAGAGGCTCCGATGGAATAATGATGATGTTGTTGTTGTTGTTGTTGTT 644  
 DB 638 CCTGAGATGAAGAGGCTCCGATGGAATAATGATGATGTTGTTGTTGTTGTTGTTGTT 697  
 QY 645 GGTGTACATTAACAGCTGTGGAATAATCTTAAGAGAGCAATCCGAATATAAGCTG 704  
 DB 698 GGTGTACATTAACAGCTGTGGAATAATCTTAAGAGAGCAATCCGAATATAAGCTG 757  
 QY 705 ATTGTGTGAACCACTGTAAGTCCAGTGTCTCAGAGAGAAAGCTGCTCACACAAG 764  
 DB 758 TATGAGTGAAGCACTGTAAGTGTCTCAGAGAGAAAGCTGCTCACACAAG 817  
 QY 765 ATTCAGGAGTGTGCTGTTTATCCCTGCTGCTTGAAGTCAATCTTCTGATGAA 824

DB 818 ATTCAGGAGTGAAGCTGTTTATACCAAGTATGTAATGTTGATCTTATTGACGAA 877  
 QY 825 GTTGTCAATATCATGATGATGAAGCAATAGAACTGCAAGACTTCTTGCGCTTAAAGAA 884  
 DB 878 GTTGTCAAGTTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 937  
 QY 885 GGCTTATTTGTGGAATATCTTCCGAGCTGCAAGCTGCTGCTTCTTCAAGTTGCAAAA 944  
 DB 938 GGTCTTCTTGTGGAATATCATCCGCTGCAAGCTGCTGCTGCAAGTTAACTTGACAG 997  
 QY 945 AGACCAAGAAATGCCCGGAAGCTTATTTGTTGCTGCTTCTTCCAGCTTCCGGAAGGTAC 1004  
 DB 998 AGCCCAAGAAACGCTGGAAGCTTATTTGTTGCTGCTTCTTCCAGCTTCCGGAAGGTAT 1057  
 QY 1005 CTGTCTCCGCTGCTTATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1064  
 DB 1058 CTATCAACGCTACTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1117  
 QY 1065 TGA 1067  
 DB 1118 TGA 1120

RESULT 3  
 AAQ44450  
 ID AAQ44450 standard; cDNA to mRNA; 1303 BP.

AC AAQ44450;  
 DT 12-OCT-1994 (first entry)  
 DB Cysteine synthase.  
 OS Cysteine synthase; plant; expression; probe; ss.  
 OS Spinacia oleracea.  
 FT Key Location/Qualifiers  
 FT CDS 52..1029 /\*tag= a  
 FT JP06038770-A.  
 PN 15-FEB-1994.  
 PD 05-FEB-1992; 92JP-00020315.  
 PF 05-FEB-1992; 92JP-00020315.  
 PR 05-FEB-1992; 92JP-00020315.  
 XX (MITS) MITSUBISHI CORP.  
 XX (MITU) MITSUBISHI KASBI CORP.  
 XX MPI; 1994-094834/12.  
 XX P-PSDB; AAR49830.  
 XX Novel gene coding cysteine synthase - used to increase the cysteine content of an agricultural plant.  
 PS Claim 2; Page 4-6; 6pp; Japanese.  
 CC The cysteine content in an agricultural product can be increased by expressing the cysteine synthase in a plant. Probes V822 and V812 used in the isolation of the gene are given in AAQ44483-84  
 CC Sequence 1303 BP; 367 A; 242 C; 311 G; 383 T; 0 U; 0 Other;  
 SQ  
 Query Match 46.7%; Score 636.4; DB 2; Length 1303;  
 Best Local Similarity 77.7%; Pred. No. 1.8e-167;  
 Matches 769; Conservative 0; Mismatches 221; Indels 0; Gaps 0;  
 QY 87 AATATGCTGTGAAGAGTCCGGAATGCTCAAGATGTTACGGAATGATGTTGTTAAACC 146  
 DB 49 AATATGCTGTGAAGAGAGCTTCTATGCTTAAAGATGATGATGATGATGATGATGATGAT 108

147 CCATAGATATCTAATAAAGTGGGATGTTGTTGCCCCGGTGTCTAAGCTG 206  
 109 CATTGGTATATCTCAACACGTCGCCGATGTTGTTGCTGTTGCTGCAAGCTG 168  
 207 GAGTTGATGAGCCATGCTCTGATGTAAGGACAGGATGGGTATAGTATGTTGAT 266  
 169 GAGGAATGGAACCTTGTCTAGGTAAAGACAGGATGGGTCACTATGATTAAGT 228  
 267 GCAGAGAGAGAGGACTTATCACTGGAAGAGTGTCTCATTTAGCCCAAGTGT 326  
 229 GCTGAAAAAAGCGGCTTATTACCTGGAGAGAGTGTCTGATTGAGCCCAAGTGA 288  
 327 AATCTGGCATTTGATTTAGCTTATGCGACGAGCCAGGGTTACAAGCTCATTAATCA 386  
 289 AATCTGGCATTTGATTTAGCTTATGCGACGAGCTAAAGTTACAAGCTCATTAATCA 348  
 387 AAGCTGCTTCTATGATGTTGAGAGAGATCATTTATTAAGCTTTGAGCTGAGTTG 446  
 349 ATGCCAGCATCAATGATGTTGAGCGGAGGACTATTCTCAGGGCTTTGTTGCTGAGCTT 408  
 447 GTTCTGACAGATCTCTGCTAAGGAAATGAAGGTGTTGAGAGAGGCTGAAGATATTG 506  
 409 ATCTTACTGATCCAGCAAAAGTATGAAGGGGCTGTTCAAGAGGCTGAGAGATCCGT 468  
 507 GCTAAGACGCCCCAATGCTCACTAATTTCAACATTTGAAAACTCTGCCAATCCAGGTT 566  
 469 GACAAAACTCTAATTCATATTAATACTACAACAGTTTGAACCTCTGCCAATCCAGGTT 528  
 567 CATTATGAAGAACCTGCTGCTCAGAGATATGAAAGGCTCCGATGGGAAATGATGATTT 626  
 529 CATTATGAAGAACCTGCTGCTCAGAGATATGAAAGGCTCCGATGGGAAATGATGATTT 588  
 627 GTTCTGAGATAGGACCTGCTGCTCAGATATCAATCAAGTGTGGAAGAAATATCTTAAAGCAG 686  
 589 GTCTCTGAATAGGAGCTGAGGATCAATACAGGTGAGGAAATATCTTAAAGCAG 648  
 687 AATCCGAATATAAGCTGATGTTGTTGGAACCAAGTGAAGTCCAGTCTCTCAGAGGA 746  
 649 AATCCGAATATAAGCTGATGTTGTTGGAACCAAGTGAAGTCCAGTCTCTCAGAGGA 708  
 747 AAGCTGTCTCACACAGATTAAGGATTTGTTGTTTATCCCTGCTGTTGGA 806  
 709 AAGCTGTCTCACACAGATTAAGGATTTGTTGTTTATCCCTGCTGTTGGA 768  
 807 GTCAATCTTCTGATGATGATGTTTCAATATCAAGTGAAGCAATAGAACTGCAAG 866  
 769 GTCAATCTTCTGATGATGATGTTTCAATATCAAGTGAAGCAATAGAACTGCAAG 828  
 867 CTCTTCTGCTTAAAGAGGCTTATTTGTTGGAATATCTTCCGAGCTGAGCTGCTGCT 926  
 829 TTGCTGCTTCAAGAGGCTTATTTGTTGGAATATCTTCCGAGCTGAGCTGCTGCT 888  
 927 GCTTTTCAAGATTCAGAAAAAGCAAGAAATGCTGGAAGCTTATTTGCTGCTTTTCC 986  
 889 GCTTTTCAAGATTCAGAAAAAGCAAGAAATGCTGGAAGCTTATTTGCTGCTTTTCC 948  
 987 AGCTTGGGAGAGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1046  
 949 AGCTTGGGAGAGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1008  
 1047 AGCATGACTTTTGAAGCTGATTTCCGCTT 1076  
 1009 AGCATGACTTTTGAAGCTGATTTCCGCTT 1038

RESULT 4  
 ADCT6317  
 ID ADCT6317 standard; DNA; 946 BP.  
 AC ADCT6317;  
 XX  
 XX 01-JAN-2004 (first entry)  
 DT

XX DNA homologous to phytopathogen resistance-related cDNA - SEQ ID 1586.  
 DE rice; yeast; poppy; plant; disease resistance; anti-fungal;  
 XX phytopathogen; gene shuffling; ds.  
 KM Unidentified.  
 OS WO2003020905-A2.  
 PN 13-MAR-2003.  
 PD 30-AUG-2002; 2002WO-US027883.  
 PF 31-AUG-2001; 2001US-0316392P.  
 PR (DOWC) DOW CHEM CO.  
 PA Shukla V, Butler H, Larrinua I, Reddy AS;  
 PI MPI; 2003-290185/28.  
 DR Novel isolated nucleic acid derived from *Nicotiana benthamiana*, *Oryza sativa*, *Saccharomyces cerevisiae*, *Trichoderma harzianum* and *Papaver rhoeas*, useful for conferring disease resistance in plants.  
 XX Claim 1; SEQ ID NO 1586; 617bp; English.  
 PS The invention relates to a novel isolated nucleic acid derived from  
 XX *Nicotiana benthamiana*, *Oryza sativa* (rice), *Saccharomyces cerevisiae*  
 CC (yeast), *Trichoderma harzianum* (Hypocrea lixii) and *Papaver rhoeas*  
 CC (poppy), and a sequence that hybridises to them under conditions of low  
 CC stringency, where expression of the nucleic acid in a plant results in a  
 CC disease resistance phenotype. The polynucleotides of the invention  
 CC demonstrate anti-fungal activity and may be useful in conferring disease  
 CC resistance in a plant against phytopathogen such as *Aspergillus flavus*,  
 CC *Gibberella fujikuroi* and *Gibberella zeae*. Furthermore, the  
 CC polynucleotides may be useful to retrieve unknown sequences and in gene  
 CC shuffling or sexual PCR procedures. The current sequence is that of the  
 CC DNA of the invention which is homologous to that of the phytopathogen  
 CC resistance-related contig cDNAs.  
 XX Sequence 946 BP; 259 A; 176 C; 250 G; 261 T; 0 U; 0 Other;  
 SQ  
 Query Match 45.4%; Score 618.2; DB 9; Length 946;  
 Best Local Similarity 78.8%; Pred. No. 2e-162;  
 Matches 737; Conservative 0; Mismatches 198; Indels 0; Gaps 0;  
 92 GCGTGTGAAGAGTCCGGAATGCGCAAGATGTTACGGAATGATGTTGTTAAACCCCAT 151  
 12 GCGGGAAGAAAGATGCAATTTGCGAAGATGTAATGATGATGATGATGATGATGATGAT 71  
 152 AGTATATCTAATAAAGTGGGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 211  
 72 GGTGACCTGATTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 131  
 212 GATGAGCCATGCTCTAGTGTGAAGAGACAGATTTGGTATAGTATGATGATGATGATGATGAT 271  
 132 CATGAGCCATGCTCTAGTGTGAAGAGATTTGGTATAGTATGATGATGATGATGATGATGAT 191  
 272 AGAGAAGGAGCTTATCAACACTGGAAGAGTGTCTCATTTAGGCCAACAAGTGTATATAC 331  
 192 GGAAGATGCGCTGATCAAACTTGGGAGAGTGTCTCATTTAGAACCTTACAGTGAACAC 251  
 332 TGGCATTTGATTAAGCTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 391  
 252 TGAAGTAGATTTGATTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 311  
 392 TGCTTCTATGAGTCTGAG 451  
 312 TTCTTCAATGAGTCTTGAAG 371  
 452 GACAGATCTGCTAAGGAGATGAAAGGTGCTGTTCAAGAGGCTGAAGAGATATTGGCTAA 511

Db	372	TACCGATCCAGCAAAAGGATGAAAGTCTATTCAGAAGGCTGAAGAGATTAAGGCCAA	431
QY	512	GAAGCCCAATGCCTACATCTTCAACAATTTGAAAACCCCTGCCAATCCCAAGGTCATTA	571
Db	432	AACACCTTAACCTCCTTATTCTTCAAGCAATTTGAAAACCCCTGCAAAACCCAAAGGTACACTA	491
QY	572	TGAAAACCACTGTGTCCAGAGATATGGAAGAAGGCTCCGATGGGAAAATTGATGCATTTGTTTC	631
Db	492	TGAGACCAACTGGTCTCTGAGATCTGSAAGAAGGCTCAAACGGGAAAGTAGATGCTCTGCTC	551
QY	632	TGGATAGGCCACTGTGTGTACATTAACAGCTGCTGGAAAATATCTTAAAGAGCAGAAATCC	691
Db	552	TGGATTTGGAACAGAGGACCAATTAACAGTTCAAGCAAGTATTTAAGAGAGCAGAAATCC	611
QY	692	GAATATTAAGCTGATTTGGTGTGAACCAAGTTGAAAAGTCCAGTGTCTCAGAGAGAAAGCC	751
Db	612	CGACATTAAGCTCTATGTGTGAACCAAGTTGAAAAGTCTATTCTTCTGAGGAAATCC	671
QY	752	TGGTCCACACAAGATTCAGAGGATTTGGTGTGTTTATCCCTGGTCTTGGAAATCAA	811
Db	672	TGGTCCGATTAAGATTCAGAGGATTTGGTGTGTTTCAATTCCTGGTCTTGGAAATCC	731
QY	812	TCTTCTTGATGAAGTGTTCAAATATCAAGTGAAGCAATAGAACAATGCAAAAGCTTCT	871
Db	732	TCTTATTGATGAAGTAAATTCAGATTTCAAGTGAAGAAAGCCATAGAACAATGCAAAAGCTTCT	791
QY	872	TGGCTTAAAGAAGGCTATTGTGGGAATATCTCCGAGCTGCAGCTGCTGCTTT	931
Db	792	GCGATTAAGAAGGTTGCTTGTGGCATATCAATCTGTGCTTCTGCTGCTGCAGCAAT	851
QY	932	TGCATTTGCAAAAAGACCAAGAAATGCCGGAAGCTTATGTGCGGTTTTTCCAGCTT	991
Db	852	CAAACTTGCTAAGCGCCCTGAAATGCTGGGAAGCTGATGTGTGTTTCCCAAGCTT	911
QY	992	CGGGAGAGCTACCTGTCTCTCCGTGCTATTGAGT	1026
Db	912	CGGGAGAGCATATCTTCTCTGTGCTCTTGAAT	946

RESULT 5  
ABZ13088  
ID ABZ13088 standard; DNA; 966 BP.  
XX  
XX ABZ13088;  
AC  
XX  
XX 21-JAN-2003 (first entry)  
DT  
XX  
XX Arabidopsis thaliana stress regulated gene SEQ ID NO 893.  
DE  
XX  
XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
KM  
XX  
XX Arabidopsis thaliana.  
OS  
XX  
XX WO200216655-A2.  
PN  
XX  
XX 28-FEB-2002.  
PD  
XX  
XX 24-AUG-2001; 2001WO-US026685.  
PF  
XX  
XX 24-AUG-2000; 2000US-0227866P.  
PR 26-JAN-2001; 2001US-0264647P.  
PR 22-JUN-2001; 2001US-0300111P.  
PR  
XX  
XX (SCRI ) SCRIPPS RES INST.  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
PA  
XX  
XX Harper JF, Kreps J, Wang X, Zhu T;  
PI  
XX  
XX WPI; 2002-304127/34.  
DR  
XX  
XX Identifying a stress condition to which a plant cell has been exposed and  
PT producing plants with increased tolerance to these abiotic stresses.

XX Claim 144; SEQ ID NO 893; 577bp + Sequence Listing; English.  
PS  
XX  
XX The invention relates to identifying a stress condition to which a plant  
CC cell has been exposed, comprising: (a) contacting nucleic acid  
CC representative of expressed polynucleotides in the plant cell with an  
CC array or probes representative of the plant cell genome; and (b)  
CC detecting a profile of expressed polynucleotides in the plant cell  
CC characteristic of a stress response. The method is useful in the  
CC production of transgenic plants, cells and seeds and in producing plants  
CC with increased tolerance to abiotic stress. The present sequence is that  
CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used  
CC in methods of the invention. Note: The sequence data for this patent is  
CC not represented in the printed specification but is based on sequence  
CC information supplied to Derwent by the European Patent Office  
XX  
SQ Sequence 966 BP; 280 A; 157 C; 261 G; 268 T; 0 U; 0 Other;

Query Match	43.7%	Score 595.2	DB 6	Length 966
Best Local Similarity	76.5%	Pred. No. 5.5e-156		
Matches 748	Conservative 0	Mismatches 218	Indels 12	Gaps 1
QY 90	ATGGCTGTGMAAGGTCGGGAATGCGCAAGATGTTACCGAATTGATTTGGTAAACCCCA	149		
DB 1	ATGGCTTCAGTGGCTCCTAAATGCTTAAAGATGTAACCTAATATAGGGAACACTCCA	60		
QY 150	TTAGTATATCTAAATAAATCTGCGGATGGTGTGTGTCGCCGGGTGCTGCTAAACTGGAG	209		
DB 61	TTAGTGTATCTAAACAAAGTAGCTTAAAGATGTGTGTGTGTAATGTTGCTGCTAAGCTTGAG	120		
QY 210	TTGATGAGAGCCATGCTCTAGTGTGAAGGACAGAGATTGGGTATAGTATGATTTGCTGATGCA	269		
DB 121	ATGATGGAACCATGTCTTAGCGCTCAAGACAGAGATTGGTTATAGTATGATTTGCTGATGCA	180		
QY 270	GAAGAGAAGGACCTTATCAACCTGGAAGAAGTGTCTCATTTGAGCCCAACAAGTGTAT	329		
DB 181	GAAGCTAAAGCACTTATTAAACCGAGAGAGTGTGTGATTGAGCCCAACAAGTGGGAAC	240		
QY 330	ACTGGCATTGGATTAGCCTTCAATGGCAGCAGCCAGGGGTACAGCTCATTAATTCAATG	389		
DB 241	ACTGAGTTGGTTAGCATTTACTGCACTGCAAGGGCTATAGCTTGTATTACCAATG	300		
QY 390	CTGTCTCTATAGACTCTTGAGAGAAGATCATTTCTATTAGCTTTTGGAGCTGAGTTGTT	449		
DB 301	CCAGCTTCATGAGTATAGAGAGAAGATCATTTCTTAGCTTTTGGAGCTGAGTTGATT	360		
QY 450	CTGACAGATCTGCTAAGGGAATGAAGGTGCTGTTCAAGAGCTGAAGAGATATTGGCT	509		
DB 361	CTTACTGATCCGGCTAAAGGTATGAAGGTGCTGTTCAAGGCGGAGAGATTTGGCG	420		
QY 510	AAGACGCCAATGCTTACATATCTCAACAATTGAAAAACCTGCCATCCCAAGGTTCAAT	569		
DB 421	AAACAACCTAATGTTATATAGCTTCAACAGTTTGAGAACTCTGCTAATCCAAAGATCCAT	480		
QY 570	TATGAACAACCTGCTGTCAGAGATATGGAAGGCTCCGATGGGAAAAATGATGCATTTGTT	629		
DB 481	TATGAGACTACTGGAACCTGAGATATGGAAGGTTCTGTGGAAGAAGTGGATGGCTTTGTT	540		
QY 630	TCTGGGATAGGCACTGTGTGTACAATAACAGGTGCTGGAATAATCTTAAAGACAGAAAT	689		
DB 541	TCTGTATTTGGTACTGTT-----GGTGTGGGAAGTATCTCAAGGAACAGAAAC	588		
QY 690	CCGAATATTAAGCTGATTGTGTGGAACAGTTGAAAGTCCAGTCTTCAGGAGGAAG	749		
DB 589	ACAAACATAAAGCTGTATGTGTGAGCCTGTTGAAGCCCTATTCTGTCCGGTGAAGAAG	648		
QY 750	CCTGTGTCCACACAAGATTCAAGGATGGTGTGCTGTTTATCCCTGTGTCTTGAAGTC	809		
DB 649	CCAGGTCCCAACAAATTCAGGTATAGGTGTGCTGTTTATTCAGGCAATTTGGATGTT	708		
QY 810	AATCTTCTTGATGAAGTGTCTCAATATCAAGTGAAGAACAATAGAACTGCAAGGCTT	869		
DB 709	GATCTTATAGATGAAGTGTCTCAGGTTCAAGTGAAGAATCCATTGACATGGCAAGGCTT	768		





DT 20-NOV-2003 (first entry)  
 XX Rice gene, SEQ ID 2735.  
 DE plant; bacterial infection; fungal infection; viral infection; rice;  
 KM gene; ds.  
 XX Oryza sativa.  
 OS WO2003000898-A1.  
 PN 03-JAN-2003.  
 PD 22-JUN-2001; 2001WO-IB001105.  
 PP 22-JUN-2001; 2001WO-IB001105.  
 PR 22-JUN-2001; 2001WO-IB001105.  
 XX (SYGN ) SYNGENTA PARTICIPATIONS AG.  
 XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
 PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
 PI WPI; 2003-175290/17.  
 DR Identifying at least one gene involved in plant resistance or response to  
 XX pathogenic infection for conferring resistance or tolerance to a plant to  
 PT bacterial, fungal or viral infection by determining or detecting plant  
 PT gene expression.  
 XX Claim 6; SEQ ID NO 2735; 899pp; English.  
 PS The present invention relates to a method (M1) for identifying genes  
 XX involved in plant resistance or response to pathogenic infection. M1  
 CC comprises identifying a gene whose expression is significantly altered in  
 CC the incompatible interaction of plant gene expression relative to  
 CC expression of the gene in an uninfected plant, in a mutant plant that  
 CC does not express a gene associated with response to pathogenic infection,  
 CC or in a corresponding incompatible or compatible interaction. (M1) is  
 CC useful for conferring resistance to resistance or tolerance to a plant to  
 CC bacterial, fungal or viral infection. The present sequence was used to  
 CC illustrate the invention.  
 CC Sequence 993 BP; 263 A; 207 C; 290 G; 233 T; 0 U; 0 Other;  
 SQ  
 QY 111 AFTGCCAAGATGTTACGGAATGATGTTGTAACCCCATTAATATCTAATAAATT 170  
 DB 40 AFTGCCAAGATGTTACGGAATGATGTTGTAACCCCATTAATATCTAATAAATT 99  
 QY 171 GCGGATGTTGTTGTTGCGGCGGTTGCTGCTAAGTGAAGTGAAGTGAAGTGAAGT 230  
 DB 100 ACGGATGTTGTTGTTGCGGCGGTTGCTGCTAAGTGAAGTGAAGTGAAGTGAAGT 159  
 QY 231 GTGAAGACAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 290  
 DB 160 GTGAAGACAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 219  
 QY 291 CTTGAAAGAGTGTCTCTCATGAGCCCAACAGTGAATGTTGTTGTTGTTGTTGTTGTT 350  
 DB 220 CTTGAAAGAGTGTCTCTCATGAGCCCAACAGTGAATGTTGTTGTTGTTGTTGTTGTT 279  
 QY 351 ATGCAAGACAGTGTCTCTCATGAGCCCAACAGTGAATGTTGTTGTTGTTGTTGTTGTT 410  
 DB 280 ATGCAAGACAGTGTCTCTCATGAGCCCAACAGTGAATGTTGTTGTTGTTGTTGTTGTT 339  
 QY 411 AGAAGATCATTTCTATGAGTTTGGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 470  
 DB 340 AGAAGATCATTTCTATGAGTTTGGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 399  
 QY 471 ATGAAGATCATTTCTATGAGTTTGGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 530

DB 400 ATGAAGATCATTTCTATGAGTTTGGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 459  
 QY 531 CTTCAACATTTGAAACCTCCCAATCCCAAGTTTATGAAACCACTGTTCCAGAG 590  
 DB 460 CTTCAACATTTGAAACCTCCCAATCCCAAGTTTATGAAACCACTGTTCCAGAG 519  
 QY 591 ATATGAAAGCTCCGATGGAATAATTGATGATTTGTTGTTGTTGTTGTTGTTGTTGTT 650  
 DB 520 ATCTGAAAGAGACAGAGGTAAGTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 579  
 QY 651 ACAATACAGTGTCTGGAATAATCTTAAAGACGAGATCCGATATATAAGCTGATGTT 710  
 DB 580 ACTATTAAGTGTCTGGAATAATCTTAAAGACGAGATCCGATATATAAGCTGATGTT 639  
 QY 711 GTGGAACCAAGTTGAAGTCCAGTCTCTCAGAGAGAAAGCTGTTCCACACAGATTCAA 770  
 DB 640 GTGGAACCAAGTTGAAGTCCAGTCTCTCAGAGAGAAAGCTGTTCCACACAGATTCAA 699  
 QY 771 GCGATGCTGCTGTTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 830  
 DB 700 GCGATGCTGCTGTTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 759  
 QY 831 CAAATATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 890  
 DB 760 CAAATATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 819  
 QY 891 TTGTTGGAATATCTTCCGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 950  
 DB 820 TTGTTGGAATATCTTCCGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 879  
 QY 951 GAAATGCGGGAAGCTTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1010  
 DB 880 GAAATGCGGGAAGCTTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 939  
 QY 1011 TCCGTTCTATTTGAGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1060  
 DB 940 TCCGTTCTATTTGAGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 989  
 RESULT 8  
 ADA70437 standard; DNA; 5259 BP.  
 ADA70437  
 ID ADA70437 (first entry)  
 DT 20-NOV-2003  
 XX Rice gene, SEQ ID 3760.  
 DE plant; bacterial infection; fungal infection; viral infection; rice;  
 KM gene; ds.  
 XX Oryza sativa.  
 OS WO2003000898-A1.  
 PN 03-JAN-2003.  
 PD 22-JUN-2001; 2001WO-IB001105.  
 PP 22-JUN-2001; 2001WO-IB001105.  
 PR 22-JUN-2001; 2001WO-IB001105.  
 XX (SYGN ) SYNGENTA PARTICIPATIONS AG.  
 XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
 PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
 PI WPI; 2003-175290/17.  
 DR Identifying at least one gene involved in plant resistance or response to  
 XX pathogenic infection for conferring resistance or tolerance to a plant to  
 PT bacterial, fungal or viral infection by determining or detecting plant





QY 67 AAGCTTGTAGTACAGATATATATGCTGTGAAAGGTCGGAAATGCCAAAGATGTTA 126  
 DB 216 AAGCTGTGTCTTTTCGCTCCATCCACATGAAAGGCTTAACATGCTGAAGATGTT 275  
 QY 127 CGGAATGATGTTAAACCCCATTAATATCTAATAAATGCGGATGTTGTTG 186  
 DB 276 CTCAGCTAATGAAAAAACCCTAATGTTATCTCAACATGTTATCAAAAGATCAGTTG 335  
 QY 187 CCCGGTGTCTAACTGAGTTGATGAGCCATGCTCTAGTGTGAAGACAGGATTTG 246  
 DB 336 CAACATGCTGCGAAGCTTGAGAGCATGAACTTGCTGAGTGTCAAGACAGGATTTG 395  
 QY 247 GGTATAGTATGATGCTGATGAGAGAGAGGACTTATCACACCTGGAAGAGTGTCC 306  
 DB 396 GCTACAGTATGATGATGATGCTGAGCAGAAAGAGTATCACACCTGGAAGAGTGTCC 455  
 QY 307 TCATTGAGCCCAACAGTGTGTAATATCTGCAATGGAATTAGCTTCATGCGACAGCCAGG 366  
 DB 456 TAGTGAGCTTACGAGTGGGAATATCTGAAATAGACTTGCTTATAGCTGTGCCAGAG 515  
 QY 367 GTTACAGCTCATTAATTAATGCTGCTCTTATAGAGTCTTGAGAGAGATCATTCTAT 426  
 DB 516 GATACAGATTACCTTGAATGACATGCTGCTGATGAGTATGAAAGAGAGTATCTTGA 575  
 QY 427 TAGCTTTGAGCTGAGTGTGCTTCTGACAGATCTGCTAAGGGAATGAAGTGTCTTC 486  
 DB 576 AAGCATTTGAGCTGAGTGTGCTTCTGACAGATCTGCTAAGGGAATGAAGAGAGTGTG 635  
 QY 487 AGAAGCTGAGAGATATTGGCTAAGACGCGCCATGCTTACATCTTCAATTTGAAA 546  
 DB 636 AGAAGCTGAGAGATTTTGAAGAAAATCTGATTTCTACATGCTTCAAGAGTGTGACA 695  
 QY 547 ACCCTGCCAATCCCAAGTTCATTATGAACCACTGCTCCAGAGATATGAAAGGCTCCG 606  
 DB 696 ATCTGCAATCCCAAGTTCATTATGAGAACCAAGTCCGAGATCTGGAAGACACAA 755  
 QY 607 ATGGAAATTTGATGATTTGTTTCTGGGATGAGCACTGCTGATCAATTAACAGTGTG 666  
 DB 756 AAGCAATGATGATTTTGTGAGGATGAGCACTGAGGAAAGATTTCTGAGATTTG 815  
 QY 667 GAAATATCTTAAAGACAGAAATCCGATATTAAGCTGATGTTGGAACCAAGTGTGAAA 726  
 DB 816 GACGTAACCTCAAGAGACGTAACCTGCTGCTGAGTATGATGATGATGATGATGATG 875  
 QY 727 GTCCAGTGTCTCAGAGAGAAAGCTGCTGCTCAGACAGATTCAGGAGTGTGCTGTT 786  
 DB 876 GCAACATACTTCTGCTGAGAAAGCTGCTGCTCAGACAGATTCAGGAGTGTGCTGTT 935  
 QY 787 TTAATCCCTGCTGCTGAGTCAATCTTCTGATGAGAGTGTGCAATATCAAGTGTG 846  
 DB 936 TTGTTCCAGCAATTAATTAATTTGGTGTGATGAGTATGAGTATGAGTATGAGTATG 995  
 QY 847 AAGCAATAGAACTGCAAGCTTCTGCTGCTTAAAGAGGCTTATTTGCGAATATCTT 906  
 DB 996 AAGCTGTAGAAATGCAAGCAATGCAATGAAGAGAGGCTTGTGTTGCAATTTTCA 1055  
 QY 907 CCGAGCTGAGCTGCTGCTGCTTTCAGATGCAAAAGACAGAAATGCGGGAAGC 966  
 DB 1056 CTGAGAGCAGCAGCTGCTGCTGCTGAGATGTTGTTAAAGACCTGAAATGCAAGAAAC 1115  
 QY 967 TTATTTGCTGCTTTTCCAGCTTGGGAGAGTACCTGCTCCGCTGATTTGAGT 1026  
 DB 1116 TTATTTGCTGCTTTTCCAGCTTGGTGTGAGAGATACCTGCTCCATTTGTTCCAGT 1175  
 QY 1027 CAGTGAGACGCAAGCTGAAAGCATGATTTTGAAGCTGAAATCCGTTAAGGCTCT 1085  
 DB 1176 CTATTCGAGAGAGTGTGAAGACATGAAGCCAGAAATGATGAGATCATTTCTTGATAT 1234

AC AAC34662;  
 XX  
 DT 17-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 7429.  
 KW Hybridisation assay; genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway; metabolic pathway;  
 KW promoter; termination sequence; ss.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-00301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 28-APR-1999; 99US-0130891P.  
 PR 30-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 30-APR-1999; 99US-0132407P.  
 PR 04-MAY-1999; 99US-0132484P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 06-MAY-1999; 99US-0132487P.  
 PR 07-MAY-1999; 99US-0132863P.  
 PR 11-MAY-1999; 99US-0134256P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138847P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139492P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.

RESULT 10  
 AAC34662  
 ID AAC34662 standard; DNA; 1503 BP.  
 XX

PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140911P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142920P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 20-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145088P.  
 PR 22-JUL-1999; 99US-0145085P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 28-JUL-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 02-AUG-1999; 99US-0146389P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.

PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 37.4%; Score 509.6; DB 3; Length 1503;  
 Best Local Similarity 70.0%; Pred. No. 6.3e-132;  
 Matches 686; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

QY 93 GCTGTTGAAGGTCGGAATTGCCAAGATGTTACGGAATTGTTGTTAAACCCCATTA 152  
 Db 360 GGTGTTGAAGGCTCAATATGCGGATACGCCGCTCAGCTTATTGGAAACTCCGATG 419  
 QY 153 GTATATCTAATTAAGTTCGGATGTTGTTGTCGCCGGTGTGCTTAACCTGAGTTG 212  
 Db 420 GTGACTTGAACATGTAAGTCAAGGCTGTGTTGCAAGTGTGCTTAAGCTTGAATC 479  
 QY 213 ATGAGCCATGCTTGAAGTGAAGACAGATGGGTATATGATGATGTCAGTCAAGAA 272  
 Db 480 ATGAACCATGTTGCAAGTGTCAAGGATAGATGGGTACAGATGATGATGCTGAA 539  
 QY 273 GAGAGGAGCTTATCAACCTGGAAGAGTGTCTTATTGAGCCACAACTGTAATCT 332  
 Db 540 GAGAAAGACTTATTAACCTGGAAGAGTGTCTTATTGGAATCTACAGTGGACACAA 599

```
OY 333 GGCATTGATAGCTTCTCATGCGAGCAGCCAGGGTTACAGCTCATTAATGAGCT 392
    |||||
DB 600 GGGATTGGCTTGCTGATTCATGCTGCTCAAAAGGCTATACTTATGAGATGCT 659
OY 393 GCTTCTATGAGCTCTGAGAGAGAAATCATCTATAGCTTTGAGCTGAGTGTCTG 452
    |||||
DB 660 GCGTCCATGAGTTTGAAGGCGGCTCTTTGAGGCGATTTGAGCTGAGCTTGTGTTA 719
OY 453 ACAGATCTGCTTAAGGAATGAAGGCTGCTTCAAGAGGCTGAAGATATTGCTAAG 512
    |||||
DB 720 ACTGAACCTGCAAAAGGTATGACTGAGCAATTCAGAGGCTGAAGAAATCTGAAAAA 779
OY 513 ACGCCCAATGCTACATACCTTCAACAATTTGAAAAACCTGCCAATCCCAAGTTTAT 572
    |||||
DB 780 ACTCCCAATTCCTACATGCTCTCCACAGTTTGAACAACCTGCCAATCCCAAGTTTAT 839
OY 573 GAAACCACTGCTCCAGAGATATGAAAGGCTCCGATGGGAAATGATGCTTTCT 632
    |||||
DB 840 GAGACGACTGCTCTGAGATTTGGGAGATACAAGAGGCAAAATGATATATTGTTGCG 899
OY 633 GGGATAGGCACTGCTGCTGACATAATACAGTCTGAAAAATATCTTAAGAGCAGATCCG 692
    |||||
DB 900 GGGATTGGAATGCTGCTGCAACTATCACTGCTGTTGTTGCTGATTCAATTAAGAAAGAAACCT 959
OY 693 AATATAAGCTGATGCTGCTGCTGCAAGCCAGTTGAAAGTCCAGTCTCTCAAGAGAAAGCCT 752
    |||||
DB 960 GAATTGAAGCTTAATGCTGCTGCAAGCCAGAAAGTCTATATCTTCTGCTGAAAGCC 1019
OY 753 GGTCCACACAGATTCAGAGGATGCTGCTTTTATCCCTGCTGCTTGAAGTCAAT 812
    |||||
DB 1020 GGACCTCACAAGATTCAGAGGATGCTGCTGATTTGTACCCCAAGAAATTTGATCTGCT 1079
OY 813 CTCTTGATGATGCTGCTCAATATCAAGTATGAGCAATAGAACTGCAAGCTTCTT 872
    |||||
DB 1080 ATTGTAGATGATATACATAGCGATTTCCAGTGAAGAGAACTATGAACTCGAAGCACTA 1139
OY 873 GCGCTTAAGAGGCTTATTTGTTGGAATATCTTCGAGCTGCACTGCTGCTTTT 932
    |||||
DB 1140 GCTCTCCAGAGGCTTGTGCTGCTGATATCTTCTGAGCTGCTGCTGCTGCAATC 1199
OY 933 CAGATTCAGAAAGACCAAGAAATGCGGAGCTTATTTGCTGCTTTTCCAGCTTC 992
    |||||
DB 1200 CAGGTTCAGAAAGACCTGAAATGCGGGAACCTCATAGCCGTTGTTCCGAGCTTC 1259
OY 993 GGGAGAGGATGCTGCTGCTGCTGCTATTTGAGTCAAGTGAAGAGCGAAGCTGAAGCAG 1052
    |||||
DB 1260 GGGAGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1319
OY 1053 ACTTTGAGCCCTGAATTC 1072
    |||||
DB 1320 CAGCCCGAGCTTGAATTTTC 1339
    |||||
```

## RESULT 11

AAC33559

ID AAC33559 standard; DNA; 1303 BP.

AAC33559;

17-OCT-2000 (first entry)

Arabidopsis thaliana DNA fragment SEQ ID NO: 3496.

Hybridisation assay; genetic mapping; gene expression control;

protein identification; signal transduction pathway; metabolic pathway;

promoter; termination sequence; ss.

Arabidopsis thaliana.

EP1033405-A2.

06-SEP-2000.

XX

```
PE 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
```



PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144684P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 23-JUL-1999; 99US-0145276P.  
PR 26-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.

PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 36.5%; Score 497.4; DB 3; Length 1303;  
Best Local Similarity 69.5%; Pred. No. 1.5e-128;  
Matches 675; Conservative 0; Mismatches 296; Indels 0; Gaps 0;

QY 90 ATGGCTGTTGAAGCTCCGGAATTGCCAAGATGTTACGGAATTGATTGTTAAACCCCA 149  
Db 93 ATGGAGAGGACCGGTGTTGCATCAAGATGATGCAACTCAATTGATAGTAACACCCCA 152  
QY 150 TTAGTATATCTAAATAAACTTGGGATGTTGTGTTCCCGGTTGCTGCTAAACTGGAG 209  
Db 153 ATGTATATCTGAACAACATTTGGATGTTGTAGTCTGTAFTGACGCTTAAGCTTGA 212  
QY 210 TTGATGAGCCATGCTCTAGTGTGAAGAGACAGATTGGTATAGTATGATGCTGATGCA 269  
Db 213 ATGATGAGCCTTCTCTAGTGTCAAGAGGCAATTGCTTATGTATGATTAAGATGCA 272  
QY 270 GAAGAGAAGGACTTATCAACCTGGAAGAGTGTCTCATTTAGGCTCAACAAGTGTAAAT 329  
Db 273 GAAGACAAGGATTGATTACTCTCGGAGAGACACATGATTGAGGCTACTCTGTGAAC 332  
QY 330 ACTGGATTGATTAGCTTCACTGAGCAGACGAGGCTTACAAAGCTCATTAATTAATG 389  
Db 333 ACCGGATTGTTAGCTTCACTGAGTCAAGTAAAGTTACAAAGTGTCTCTCAATG 392  
QY 390 CCTGCTTCTATGAGTCTTGAAGAGAAATCATTTAGCTTTGGAGTGAAGTTGCTT 449  
Db 393 CCTCATCATGAGCTTGAAGAGAAATCATTTAGCTTTAGCTTTAGCTGAGGTTGCTT 452  
QY 450 CTGACAGATCTGCTTAAGGAAATGAAGGTGCTTCAAGAGGCTGAAGAGATTTGGCT 509  
Db 453 CTCACAGATCTCTAGTAAGGCGCTTCAAGGATTAATGACAAAGCTGAAGAGATTTAGC 512  
QY 510 AAGACGCCCAATGCTCTACATATCTCAACAATTGTAACCCCTGCCAATCCCAAGTTGCT 569  
Db 513 AAACTCCAGATGATGATGCTGTAAGAACAGTTCAAAAATCTTCAAACCCGCAACTCAT 572

QY	570	TATGAAMCCACTGGTCCAGACAGATATGGAAGGAGGCTCCGATGGGAAAATTGATGCATTTGTT	629
Db	573	TATGAAACCAACGGGTCCAGAGATATGAGAGACTCTGCAGGGGAAGTAGACATATTGGTT	632
QY	630	TCTGGATAGGCACCTGGTGGTCAATATACAGGTGCTGAAAATATCTTAAGACAGCAAT	689
Db	633	GCCGGTGTGGAACCTGGTGGAAACGCTTTCCGATCAGGAAGATTCTCAAGGAGAAAT	692
QY	690	CCGAATATTAAGCTGATTGGTGTGGAACCAAGTTGAAGTCCAGTGCCTCAGAGAGAAAAG	749
Db	693	AAAGACTTTAAGTTTATGCTGTGGAACCTACAGAAAGTCCGGTAATTAAGTGAGGCCAA	752
QY	750	CCTGTGCCACACAGATTCAAGGGAATGGTGTGCTGTTTATCCCTGCTGTCTGGAAGTC	809
Db	753	CCGGGTACACATTGTGATCCAGGTAFTGGGCTGACTCATCCAGACAATTGGATTTC	812
QY	810	AATCTTCTTGATGAAGTGTTCAAATATCAAGTGATGAAGCAATAGAACTGCAAGCTT	869
Db	813	AACGTTCTTGATGAAGTCAATCCAAAGTGACAAGTGTGGAAGCAATTGAAACAGCCAACTT	872
QY	870	CTTGCGCTTAAAGAAGGCTTATTTGTGGGAATATCTCCGAGCTGCAGCTGCTGCTGCT	929
Db	873	CTTGCCCTGAAGAAGGATTACTGTTGGGAATATCTTCTGAGCTGCTGCAGCCGCTGCG	932
QY	930	TTTCAGATTGCCAAAAAGACAGAAAATGCCGGGAAGCTTATGTGCGCTTTTCCACAG	989
Db	933	ATTAAGGTGGCAAGCGGCCAGAAAACGCCGCAACTCATAGTGTGATTTCCCTTAGC	992
QY	990	TTCCGGGAGAGGTACTGTCTCCGTCCGTCTATTGAGTGAAGAGACGCGAAGCTGAAGC	1049
Db	993	GGTGGGAACGTTACTTATCGACCTCATTTGTTGAATCAGTCAGACATGAAGCAGAGAAT	1052
QY	1050	ATGACTTTTGA	1060
Db	1053	TTGCCAATTCA	1063

```

RESULT 12
ADCC64221
ID ADCC64221 standard; cDNA; 975 BP.
XX
AC ADCC64221;
XX
DT 18-DEC-2003 (first entry)
XX
DE Thalecress cysteine synthase coding region.
XX
XX Thalecress; cysteine synthase; ss; O-acetyl-L-
KM coumarin dye; cysteine-coumarin conjugate; he
KM bactericide; insecticide; plant.
XX
OS Arabidopsis thaliana.
XX
XX
FH Key Location/Qualifiers
CDS 1..975
FT /tag= a
FT /product= "Cysteine synthase"
FT /EC_number= "3C4.2.99.8"
FT

```

PD 08-MAY-2003.  
XX  
PF 13-JUL-2001; 2001US-00905290.  
XX  
PR 13-JUL-2001; 2001US-00905290.  
XX  
PA (RICE/) RICE J.  
PA (LANN/) LANNING B.  
PA (CRAW/) CRAWFORD J.  
PA (NYEG/) NYE G.  
XX  
PI Rice J, Lanning B, Crawford J, Nye G.

XX WPI; 2003-777218/73.  
DR P-PSDB; AD064218.  
DR  
XX  
PT Determining cysteine synthase activity by contacting cysteine produced by  
PT the enzyme in a reaction mixture with a coumarin dye, exciting the formed  
PT conjugate and detecting the fluorescent light emitted by conjugate.  
PT  
XX  
PS Disclosure; SEQ ID NO 3; 32pp; English.

The invention relates to determining (M1) cysteine synthase activity, by combining O-acetyl-L-serine, sulphide and cysteine synthase to form reaction mixture under conditions suitable for production of cysteine, contacting reaction mixture with coumarin dye capable of forming cysteine-coumarin conjugate, and subjecting the reaction mixture to ultraviolet (UV) light to excite the conjugate, and detecting fluorescent light emitted by the conjugate. The cysteine synthase is a plant cysteine synthase appearing as ADC64218. The method is useful for determining cysteine synthase activity by quantitating the cysteine formed by the enzyme and for identifying a test compound as a herbicide candidate (which involves performing the method in the presence and absence of the compound, and comparing the amount of the fluorescent light emission in the presence and absence of the test compound, where a decrease in the amount of the fluorescent light emission in the presence of the test compound indicates that the test compound is a herbicide candidate. The method is also useful for identifying compounds capable of selectively inhibiting plant, fungal and/or bacterial cysteine synthase activity (which involves performing the method in the presence of plant cysteine synthase and determining the inhibitory activity of the compound based on fluorescent light emission, repeating the steps using a fungal or bacterial cysteine synthase, and identifying compounds that selectively inhibit plant, fungal or bacterial cysteine synthase). The method is useful for identifying a compound as a candidate for herbicides, fungicides, bactericides and insecticides. The present sequence is the A. thaliana cysteine synthase coding region.

Sequence	975 BP; 297 A; 187 C; 253 G; 238 T; 0 U; 0 Other;
Query Match	36.4%; Score 495.8; DB 9; length 975;
Best local Similarity	69.4%; Pred. No. 3.8e-128;
Matches 674; Conservative	0; Mismatches 297; Indels 0; Gaps 0;

QY	90	ATGGCTGTGAAAGGTCGGGAATTGCCAAGATGTTACGGAAATTGATTGGTAAACC	149
Db	1	ATGAGAGGAGACCGGTGTTGATCAAGGATGATGCAACTCAATTGATAGGTACACCCCA	60
QY	150	TTAGTATATCTAAATAAACTGCGGATGGTGTGTGTGTCGCCGGGTGCTGCTAACTGGAG	209
Db	61	ATGTATATCTGAACAACAATTGTGATGGTTGTGTAGCTCGTATTGCAGCTAAGCTTGA	120
QY	210	TTGATGAGCCATGCTCTGATGTGAAGACAGAGATTGGGTATAGTATGATTGCTGATGCA	269
Db	121	ATGATGAGACCTGTGCTCTAGTGTCAAGAGCGGAATGCTTATGTGTATGATTAAGATGCA	180
QY	270	GAAAGAGAAGGACTTATCACACTGGAAAGAGTGTCTCATTTAGGCCCAACAAGTGTAT	329
Db	181	GAGACAAAGGATTGATTACTCTGGAGAGAGCACACTGATTTGAGGCTTACCTGTGTAAC	240
QY	330	ACTGCCATTGATTAGCCTTCATGGCAGCGCAGGGTTACAAGCTCATTAATTACAATG	389
Db	241	ACCGGATTGTTAGCCTTCATGGGTGCAGCTAAAGGTTACAAGTGTCTCACAATG	300
QY	390	CCTGCTTCTATGAGTCTTGAAGAGAATCATTTCTATTAGCTTTTGAGCTGAGTTGTT	449
Db	301	CCCTCATCATATGAGCCTTGAGAGAAAATCATTTCTTTAGCAATTAGGTGCGAGTTCAAC	360
QY	450	CTGCACAGATCCTGCTAAGGAATGAAGGTGCTGTTCAGAAAGCTGAAGAGATATGGCT	509
Db	361	CTCACAGATCTCTAGTAAAGCGCTCAAGGAATAATCGACAAAGCTGAAGAGATATGTAGC	420
QY	510	AAGACGCCCAATGCTTACACTTCAACAATTGAAAAACCTGCGCAATCCCAAGTTCAAT	569
Db	421	AAAAATCCAGATAGTATCATGTCTAGAACAGTTCAAAAATCCTTCAAAACCGCAAACTCAT	480

QY 570 TATGAAACCACTGTTCCAGAGATATGAAAGGCTCCGATGGAAAATGATGATTTGTT 629  
 DB 481 TATCGAACCAAGGTTCCAGAGATATGAGAGACTCTGCAAGGGAAGTAGACATATTTGTT 540  
 QY 630 TCTGGGATAGGCACTGTTGTTCAATTAACAGGTGCTGGAATAATCTTAAGAGCAAT 689  
 DB 541 GCCGGTGTGGAAGTGTGGAAGCTTTCCGATCAGGAAGATTCTCAAGGAGAGAT 600  
 QY 690 CCGAATATTAAGCTGATTTGTTGGAACCAAGTTGAAGTCCAGTCTCTCAGAGAGAA 749  
 DB 601 AAAGACTTTAAGTTTATGTTGTTGGAACCTACAGAAAGTCCGTTAATAGTGAGGCA 660  
 QY 750 CCTGCTCCACACAGATTCAAGGATGCTGCTGTTTATCCCTGCTGCTTGAAGTC 809  
 DB 661 CCGGTACACATTGATTTCAAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
 QY 810 AATCTTCTGATGAGTTGTTCAATTAATCAAGTGAAGCAATAGAACTGCAAGCTT 869  
 DB 721 AAGCTTCTGATGAGTCAATCAAGTGAAGCAATAGAACTGCAAGCTT 780  
 QY 870 CTGCGCTTAAGAAGGCTATTTGGAATATCTTCCGAGCTGCTGCTGCTGCT 929  
 DB 781 CTGCGCTTAAGAAGGCTATTTGGAATATCTTCCGAGCTGCTGCTGCTGCTGCT 840  
 QY 930 TTTCAGATTGCAAAAAGACCAAGAAATCCGGAAGCTTATTTGCTTTTCCAGC 989  
 DB 841 ATAAAGTGGCAAGCGCCAGAAACCGCGCAACTCATAGTTGATTTTCCCTAGC 900  
 QY 990 TTGCGGAGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1049  
 DB 901 GGTGGGAGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
 QY 1050 ATGACTTTTGA 1060  
 DB 961 TTGCCAATTC 971

RESULT 13  
 AAC44727  
 ID AAC44727 standard; DNA; 1200 BP.  
 XX AAC44727;  
 AC  
 XX  
 DT 18-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 43907.  
 XX  
 KW Hybridisation assay; genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway; metabolic pathway;  
 KW promoter; termination sequence; ss.  
 XX Arabidopsis thaliana.  
 OS  
 XX  
 PN EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-00301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.  
 PR 28-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 30-APR-1999; 99US-0132407P.  
 PR 04-MAY-1999; 99US-0132484P.  
 PR 06-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 06-MAY-1999; 99US-0132487P.  
 PR 07-MAY-1999; 99US-0132863P.  
 PR 11-MAY-1999; 99US-0134256P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138847P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139492P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140991P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142920P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.



PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 23-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.

PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 36.4%; Score 495.8; DB 3; Length 1200;  
Best Local Similarity 69.4%; Pred. No. 4.2e-128;  
Matches 674; Conservative 0; Mismatches 297; Indels 0; Gaps 0;

OY 90 ATGCTGTGAAAGTCCGGAATGCCAAGATGTACGAATGATGTAACCCCA 149  
DB 226 ATGAGAGGACCCGGTGTGATCAAGATGATGCACTCAATGATAGTAAACCCCA 285  
OY 150 TTAGTATATCTAATAACTTGGGATGTTGTGTTGCCGGTGTGCTAACTGGAG 209  
DB 286 ATGATATATCTGAACAACATTTGTGATGTTGTGTAGCTCGATTTGACGTAAGTGA 345  
OY 210 TTGATGAGCCATGCTCTAGTGTGAAGACAGAGTTGGTATGATGCTGATGCA 269  
DB 346 ATGATGAGCCTTGTCTAGTGTCAAGAGCGGAATTGCTTATGATTAAGATGCA 405  
OY 270 GAAGAGAGGACCTTATCACACCTTGAAAGAGTGTCTCATTTGACCAACAGTGTAT 329  
DB 406 GAAGACAGGATGATTAATCTCTCTGGAAGAGACACACTGATTTAGGCTAAGTGA 465  
OY 330 ACTGGCATTTGATTAAGCTTCAATGAGACAGCCAGGGTTACAACTCATATTAATG 389  
DB 466 ACCGGATGTTTGAAGCTTCAATGAGACAGCTTAAGGTTACAACTGCTCAATG 525  
OY 390 CCTGCTTATGAGTCTTGAAGAGAAATCAATTTATTTAGCTTTGAGCTGATTTG 449  
DB 526 CCTCATCATGAGCCTTGAGAGAAATCAATTTTATGATTTAGCTGAGGCTTCA 585  
OY 450 CTGACAGATCTGCTTAAGGAAATGAAGTGTGTTCAAGAGCTGAAGATTTGGCT 509  
DB 586 CTACAGATCTTAAGGAGGCTTCAAGAAATTAATGACAAAGCTGAAGATATGTAGC 645  
OY 510 AAGAGCCCAATGCTTACATACATTTCAAAATTTGAAAACCTGCAATCCCAAGTTCA 569  
DB 646 AAAAATCCAGATAGTATCATGCTGTAAGCAATTTCAAAAATCTTCAAACTGCAATCAT 705  
OY 570 TATGAAACCACTGCTCAGAGATATGAAAGGCTCCGATGGAATAATGATGATTTGTT 629  
DB 706 TATGAAACCACTGCTCAGAGATATGAAAGGCTCAGAGGAGTGAACATATTTGTT 765  
OY 630 TCTGGATAGGCACTGCTGATCAATTAAGAGTGTGGAATAATTTAAAGAGCAAT 689  
DB 766 GCGGTGTGGAATGCTGGAAGCTTTCCGATCAGAAAGTCTTCAAGAGAAAGAT 825  
OY 690 CCGAATATAAGCTGATTTGTTGGAACCAAGTGAAGTCCAGTGTCTCAGAGGAAAG 749  
DB 826 AAAGACTTAAAGTTATGTTGTAACCTTACAGAAAGTCCGTAATTAAGTGAAGGCA 885  
OY 750 CTTGCTCACAACAAGTTCAAGGAGTGTGCTGTTTATCCCTGCTGCTTGAAGTC 809  
DB 886 CCGGTACACATTTGATCCAAAGTATTTGGGCTGACTCATCCAGACAAATTTGATTT 945







QY 1049 CATGACTTTGA 1060  
 Db 1051 TTGGCAATTTCA 1062

RESULT 15  
 ADD17591  
 ID ADD17591 standard; DNA; 725 BP.

AC ADD17591;  
 XX 15-JAN-2004 (first entry)

XX DNA (SeqID 1659) that confers an altered visual phenotype in plants.  
 DE ds; visual phenotype; plant; architecture; leaf surface; chlorotic;  
 KM bleaching; etching; wet leaf; stunting; elongation; texture;  
 KM agronomic trait; growth regulation; dwarf variety; insect resistance;  
 KM heat stress; transgenic.

XX Unidentified.

XX WO2003020741-A1.

XX 13-MAR-2003.

XX 30-AUG-2002; 2002WO-US027880.

XX 31-AUG-2001; 2001US-0316326P.

XX (DOWC ) DOW CHEM CO.  
 PA (DOWC ) DOW AGROSCIENCES LLC.

XX Crosley R, Skokut T, Ruegger M, Larrinua I, Shukla V;

XX WPI; 2003-300858/29.

XX Novel isolated nucleic acid derived from *Nicotiana benthamiana*, *Oryzae*  
 PT sativa, *Saccharomyces cerevisiae*, and *Papaver rhoeas*, useful for  
 PT conferring altered visual phenotypes in plants.

PS Claim 1; SEQ ID NO 1659; 517bp; English.

XX This invention relates to the identification and isolation of novel  
 CC nucleic acid molecules that confer altered visual phenotypes in plants.  
 CC Specifically, it refers to modifications of plant architecture and/or  
 CC leaf surface features in plants, such as chlorotic, bleaching, etching,  
 CC wet leaf, stunting, elongation and texture phenotypes, which are thought  
 CC will be agronomic traits beneficial to the farmer. As such, these novel  
 CC phenotypes can affect growth regulation i.e. useful for creating dwarf  
 CC varieties, exhibit resistance to insects or heat stress, confer changes  
 CC in pigment content such that plants have enhanced vitamin production or  
 CC delayed senescence and also for example produce plants that control the  
 CC production of ethylene. Furthermore, the present invention comprises  
 CC generating transgenic plants, as well as reproducibly altering the visual  
 CC phenotype of plant seeds, plant tissues and plant cells containing the  
 CC polynucleotides described herein. This polynucleotide is a homologue of a  
 CC DNA sequence that confers an altered visual phenotype when expressed in  
 CC plants, the method of the invention.

XX Sequence 725 BP; 206 A; 133 C; 195 G; 191 T; 0 U; 0 Other;

Query Match 35.4%; Score 482; DB 9; Length 725;  
 Best Local Similarity 79.7%; Pred. No. 2.4e-124;  
 Matches 569; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 92 GGCTGTGAAGCTCCGGAATGCCAAGATGTTACGAATTGATGTAAGAACCCCAT 151  
 Db 12 GCGCGGAGAAAGATGGAATTCGGAAGATGTAAGTGAATGATCGGTAACACTCCTTT 71

QY 152 AGTATATCTAAATAAATCTCGGATGCTGTGTTGCTCCCGGTTGCTGCTAAACTGAGTT 211

Db 72 GGTGTACTGAATAATGTTGTGATGGGTGTGTTGCTGCGCTTGTCTGCCAAGCTCGAAG 131  
 QY 212 GATGAGCCATGCTCTAGTGTGAAGACAGATTTGGTATAGTATGTTGCTGATGAGA 271  
 Db 132 CATGAGCCATGCTCTAGTGTGAAGATGATTTGTTATAGTATGATGATGATGCTGA 191  
 QY 272 AGAGAAAGGACTTATACACCTGGAAGAGTGTCTCATTTGAGCCCAAGTGTATATAC 331  
 Db 192 GGAGATGCTGATCAAACTGGCGAGAGTGTCTCATTTGAACTTACAAAGTGAACAC 251  
 QY 332 TGGCATTTGATTAAGCTTATGAGCAGCAGGAGGTTACAAAGCTCATTAATTAATGCC 391  
 Db 252 TGGAGTAGGATTTGATTTATGCTGCTGTGTTAAAGATACAAACTCATTAACGATGCC 311  
 QY 392 TGCTTCTATGATTTGAGAGAGAAATCATTTATTAAGCTTTTGGAGCTGATGTTGCT 451  
 Db 312 TTCTTCAATGATTTGAGAGAGAAATTTATTTGCTGCTTTTGGTGTGATGTTGCT 371  
 QY 452 GACAGATCTGCTAAGGAATGAAGGTGCTGTTCAAGAGCTGAAGATATGCTAA 511  
 Db 372 TACTGATCCAGCAAAAGGATGAAGGTTCTATTCAGAGGCTGAAGAGATTAAGGCCAA 431  
 QY 512 GACGCCAATGCTTACATTAATTTCAACAAATTTGAACCTTCCCAATCCCAAGTTCAATTA 571  
 Db 432 AACACCTAATCTCTTATTTCTTCAAGCAATTTGAACCTTCAATCCAAAGTACACTA 491  
 QY 572 TGAACCACTGTTCCAGATATGGAAGGCTCCGATGGGAATTTGATGATTTGTTTC 631  
 Db 492 TGAACCACTGTTCTGAGATCTGGAAGGCTCAAGGGAAGTATGCTCTAGTCTC 551  
 QY 632 TGGATAGGCACTGTGTGTAATTAACAGGTGCTGGAATATATCTTAAAGCAGAAATCC 691  
 Db 552 TGAATTTGAACAGAGGCAATTAACAGGTTCAGGCAAGTATTTAAGAGCAGAAACC 611  
 QY 692 GATATTAAGCTGATTTGTGTGGAACCAAGTTGAAGTCCAGTCTCTCAGAGGAAAGCC 751  
 Db 612 CGACATTAAGCTTATGTGTGTGAACCAAGTTGAAGTCTTATTTCTTGAAGAAATCC 671  
 QY 752 TGGTCCACACAAGATTTCAAGGATTTGCTGCTGTTTATCCCTGCTGTTGGA 805  
 Db 672 TGGTCCGATTAAGATTTCAAGGATTTGCTGCTGTTTATCCCTGCTGTTGGA 725

Search completed: May 21, 2004, 17:53:44  
 Job time : 659 secs

Mon May 24 08:18:50 2004

us-09-931-457a-30.fst

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2004, 11:41:31 ; Search time 3959 Seconds  
(without alignments)  
10273.375 Million cell updates/sec

Title: US-09-931-457A-30  
Perfect score: 1362  
Sequence: 1 actttgtatgtcgtatagatag.....aaaaaaaaaaaaaaaaaaaa 1362

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	768.8	56.4	772	14	CAB01406 sa05c03.
2	758	55.7	771	12	BI970016 BI970016
3	754.4	55.4	776	12	BI969937 GM830009B
4	692.4	50.8	719	10	BE658125 BE658125 GM700004B

5	651.4	47.8	803	12	BG648907	BG648907	EST510526
6	642	47.1	678	10	AW311624	AW311624	sg43b07.Y
7	636.4	46.7	929	14	CF513886	Cabud0007	CF513886 Cabud0007
8	630.4	46.3	632	14	CA801673	sati6d04.	CA801673 sati6d04.
9	629.2	46.2	910	14	CB974516	CAB30005	CB974516 CAB30005
10	627.2	46.0	724	10	BF003458	EST431956	BF003458 EST431956
11	626.2	46.0	791	12	BG595822	EST494500	BG595822 EST494500
12	625.4	45.9	994	14	BE660317	EST706089	BE660317 2-D7 Gmax
13	625	45.9	653	10	BE660317	EST717288	BE660317 2-D7 Gmax
14	612.8	45.0	967	14	CK271210	saq07c05.	CK271210 EST712232
15	606	44.5	613	13	BQ630428	EST489667	BQ630428 saq07c05.
16	599	44.0	937	14	CK266154	EST510489	CK266154 EST712232
17	595.6	43.7	785	12	BG587892	EST593689	BG587892 EST489667
18	595.4	43.7	734	12	BG648870	EST593689	BG648870 EST510489
19	588.2	43.2	673	12	BM815595	EST593689	BM815595 EST593689
20	584.6	42.9	920	14	CK259857	EST705935	CK259857 EST705935
21	581.6	42.7	907	14	CK260012	EST706090	CK260012 EST706090
22	581.6	42.7	917	14	CK259858	EST725584	CK259858 EST705936
23	578.4	42.5	934	29	CG848066	ZMMBB031	CG848066 ZMMBB031
24	577.2	42.4	935	14	CK279506	EST705567	CK279506 EST705567
25	576.2	42.3	881	14	CK259489	EST530680	CK259489 EST705567
26	573.4	42.1	655	12	BI309270	EST530680	BI309270 NF100C02D
27	572.2	42.0	653	12	BG451772	EST705566	BG451772 NF100C02D
28	571.8	42.0	873	14	CK259488	EST705566	CK259488 EST705566
29	570.6	41.9	845	14	CF200608	RR890915N	CF200608 RR890915N
30	570.4	41.9	658	12	BI264457	NF113H01P	BI264457 NF113H01P
31	568.6	41.7	802	14	CB974696	CAB30005	CB974696 CAB30005
32	567.4	41.7	570	12	BM891119	sam23b05.	BM891119 sam23b05.
33	566	41.6	745	12	BM815488	EST593582	BM815488 EST593582
34	566	41.6	886	14	CK265667	EST711745	CK265667 EST711745
35	562.2	41.3	792	14	CF512887	CABud0005	CF512887 CABud0005
36	561.2	41.2	813	14	CF205785	RR890915I	CF205785 RR890915I
37	559.2	41.1	686	12	BI264323	NP118G04P	BI264323 NP118G04P
38	557.4	40.9	1483	11	AY103651	Zea mays	AY103651 Zea mays
39	550.8	40.4	554	12	BM891219	sam24d01.	BM891219 sam24d01.
40	546.6	40.1	709	14	CA411511	B813_F-P	CA411511 B813_F-P
41	544.8	40.0	581	10	AM830316	sm25d07.Y	AM830316 sm25d07.Y
42	544.6	40.0	623	10	AW775920	EST334985	AW775920 EST334985
43	543.8	39.9	627	10	BF005244	EST433742	BF005244 EST433742
44	543.4	39.9	856	14	CF652605	65-L02057	CF652605 65-L02057
45	533	39.1	666	12	BI265217	NF103E11I	BI265217 NF103E11I

ALIGNMENTS

RESULT 1  
CAB01406  
LOCUS  
DEFINITION  
sa05c03.y2 Gm-cl062 Glycine max cDNA clone SOYBEAN CLONE ID:  
Gm-cl062-7397 5' similar to SW:CYSK\_CITLA Q43317 CYSTEINE SYNTHASE  
ACCESSION  
CAB01406  
VERSION  
CAB01406.1 GI:26058492  
KEYWORDS  
EST.  
SOURCE  
Glycine max (soybean)  
ORGANISM  
Glycine max  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
REFERENCE  
1 (bases 1 to 772)  
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corryell, V.,  
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,  
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,  
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,  
Schurk, R., Ritter, B., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,  
McCann, R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine

Mon May 24 08:18:50 2004

us-09-931-457a-30.rst

**Page 2**

44444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)  
This clone is available through: ResGen, Invitrogen Corp. 2130  
South Memorial Parkway Huntsville, AL 35801 For further information  
call: (800)-533-4363 or contact: [ccu@resgen.com](mailto:ccu@resgen.com) web site:  
[www.resgen.com](http://www.resgen.com)

## FEATURES

**Source**

```

/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl062-7397"
/tissue_type="stem tissue of greenhouse grown plants"
/dev_stage="1 month old"
/lab_host="DH10B"
/clone_lib="Gm-cl062"
/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI. The cDNA library was constructed from mRNA isolated from stem tissue of 1 month old greenhouse grown plants for the cultivar Raiden. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

```

## ORIGIN

Query Match	56.4%	Score 768.8	DB 14	Length 772	
Best Local Similarity	99.7%	Pred. No. 2.2e-130			
Matches 770	Conservative	0	Mismatches 2	Indels 0	Gaps 0
QY 59	CTCAAGTCAAGCTTTGTAGTGAGCAGATATATGCTGTTGAAAAGTCCCGAATTGCCAA	118			
DB 1	CTCAAGTCAAGCTTTGTAGTGAGCAGATATATGCTGTTGAAAAGTCCCGAATTGCCAA	60			
QY 119	AGATGTTACGGAATTGATGTTAAACCCTTAGTATATCTAAATAACTTGCCGATGG	178			
DB 61	AGATGTTACGGAATTGATGTTAAACCCTTAGTATATCTAAATAACTTGCCGATGG	120			
QY 179	TTGTGTGCCCCGGTTGCTGCTTAACTGGAGTTGATGAGCCATGCTCTAGTGTGAAGA	238			
DB 121	TTGTGTGCCCCGGTTGCTGCTTAACTGGAGTTGATGAGCCATGCTCTAGTGTGAAGA	180			
QY 239	CAGGATTGGGTATAGTAATGATTGCTGATGCAGAGAAGAGGACTTATCACACCTGGAA	298			
DB 181	CAGGATTGGGTATAGTAATGATTGCTGATGCAGAGAAGAGGACTTATCACACCTGGAA	240			
QY 299	GAGTGTCTCATTGAGCCACAAGTGTATATCTGGCAATTGATTGACCTTCATGGCAGC	358			
DB 241	GAGTGTCTCATTGAGCCACAAGTGTATATCTGGCAATTGATTGACCTTCATGGCAGC	300			
QY 359	AGCCAGGGGTTACAAGCTCATTAATTACAATGCGCTGCTTATGAGTCTTGAGAGAGAAAT	418			
DB 301	AGCCAGGGGTTACAAGCTCATTAATTACAATGCGCTGCTTATGAGTCTTGAGAGAGAAAT	360			
QY 419	CATTCTATTAGCTTTTGAAGCTGAGTTGGTTCTGACACAGTCTGCTAAGGGAATGAAGG	478			
DB 361	CATTCTATTAGCTTTTGAAGCTGAGTTGGTTCTGACACAGTCTGCTAAGGGAATGAAGG	420			
QY 479	TGCTGTTCAGAAAGGCTGAAGAGATATTGGCTAAGACGCCCCAATGCTACATACCTTCAACA	538			
DB 421	TGCTGTTCAGAAAGGCTGAAGAGATATTGGCTAAGACGCCCCAATGCTACATACCTTCAACA	480			
QY 539	ATTGAAAACCCCTGCCAATCCCAAGGTTCAATTATGAAAACCACTGGTCCAGAGATATGAA	598			
DB 481	ATTGAAAACCCCTGCCAATCCCAAGGTTCAATTATGAAAACCACTGGTCCAGAGATATGAA	540			

OY	599	AGGCTCCGATGGGAAAATTGATGCATTTGTTCTTCTGGGATAGGCACTGTGGGTACAATAAC	658
Db	541	AGGCTCCGATGGGAAAATTGATGCATTTGTTCTTCTGGGATAGGCACTGTGGGTACAATAAC	600
OY	659	AGGTGCTGGAATAATATCTTAAAGAGCAGAATCCGAATATATAAGCTGATTGGTGTGGAAC	718
Db	601	AGGTGCTGGAATAATATCTTAAAGAGCAGAATCCGAATATATAAGCTGATTGGTGTGGAAC	660
OY	719	AGTTGAAAGTCCAGTGTCTTCAGAGAGAAAGCCTGTGTCACACAAGAATTCAGGGGATTGG	778
Db	661	AGTTGAAAGTCCAGTGTCTTCAGAGAGAAAGCCTGTGTCACACAAGAATTCAGGGGATTGG	720
OY	779	TGCTGCTTTATCCCTGGTGTCTTGGAAAGTCAATCTTCTTGATGAAGTTGTT	830
Db	721	TGCTGCTTTATCCCTGGTGTCTTGGAAAGTCAATCTTCTTGATGAAGTTATT	772

	771 bp	mRNA	linear	EST 23-OCT-2001
RESULT 2				
B1970016/c				
LOCUS	B1970016			
DEFINITION	GM830009B12A07 Gm-r1063 Glycine max cDNA clone Gm-r1083-3109 3'			
	mRNA sequence.			

ACCESSION	BI970016	
VERSION	BI970016.1	GI:16344421
KEYWORDS	EST.	
SOURCE	Glycine max (soybean)	
ORGANISM	Glycine max	

REFERENCE AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 771)	Vodkin, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Corvett, V., Erpelind, J., Radh, C., Shoop, E., Pardinas, J., Liu, L. and Lewin, H.	A Functional Genomics Program for Soybean (NSF 9872565)	Unpublished (1999)
Other_ESTS: AW830316	corresponding to Gm-cl028-4934 (5')		

Contact: Vodka, L.O., PI, A Functional Genomics Program for Soybean (NSF 9872565)  
Lewin, H. A., Director, Keck Center for Comparative and Functional Genomics  
University of Illinois  
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA  
Tel: (217) 244-6147  
Fax: (217) 333-4582  
Email: l-vodka@uiuc.edu  
This clone is available through: Incyte Genomics, 4633 World Parkway Circle St. Louis, Missouri 63134. Phone (800) 430-0030 or (314) 427-3222 FAX: (314) 427-3324. Web site: <http://www.incyte.com/reagents/catalog.jsp?page=clones/collaboration/index>  
Seq primer: 5'-TTTTTTTTTTTTTTTTT(A/C/G)-3'.

## FEATURES

### Source

```

/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="Gm-r1083-3109"
/clone_lib="Gm-r1083"
/note="The library Gm-r1083 is a sequence-driven, reracked
set of 4,992 clones selected from cDNA libraries from
various tissues and stages of development of soybean. It
represents 1117 sequences from the progenitor library
Gm-cl009 (from mature roots of 2 month old greenhouse
grown 'Williams' soybean plants); 820 sequences from the
progenitor library Gm-cl013 (from 2 to 3 week old whole
plants of Williams); and 3055 sequences from library
Gm-cl028 (from 'Supermod' plants whose seedlings were
inoculated with Bradyrhizobium japonicum, courtesy of Dr.
Gary Stacey). The 5' ESTs of the source clones from the
different progenitor libraries was used to select
singletons, or a representative of each contig, which were

```



reracked to form library Gm-r1083. The cDNA clones of the reracked Gm-r1083 library were then sequenced at the 3' end. The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Center for Computational Genomics and Bioinformatics, University of Minnesota, <http://web.abc.umn.edu/biodata/nfsioy/>. Reracking was performed by Incyte Genomics, St. Louis, <http://www.incyte.com>, and 3' sequencing by the Keck Center for Comparative and Functional Genomics, University of Illinois, <http://www.lie.uic.edu/biotech/keck.html>. Note: The corresponding 5' EST from each clone in the Gm-r1083 library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the Incyte Genomics clone ID of the original cDNA library that is also listed under 'OTHER EST'."

# ORIGIN

Query Match 55.7%; Score 758; DB 12; Length 771;  
Best Local Similarity 98.3%; Pred. No. 2e-128;  
Matches 758; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

462 GCTAAGGAATGAAAGGCTGTTGAGAGGCTGAAGAGATTTGGCTAAGACGCCCAT  
771 GCTAAGGAATGAAAGGCTGTTGAGAGGCTGAAGAGATTTNNNTAAGACGNNNAT  
522 GCTAAGGAATGAAAGGCTGTTGAGAGGCTGAAGAGATTTGGCTAAGACGCCCAT  
711 GCTAAGGAATGAAAGGCTGTTGAGAGGCTGAAGAGATTTNNNTAAGACGNNNAT  
582 GGTCCAGAGATATGAAAGGCTCCGATGGGAAATTTGATGATTTGTTCTGGATAGG  
651 GGTCCAGAGATATGAAAGGCTCCGATGGGAAATTTGATGATTTGTTCTGGATAGG  
642 ACTGGGTGTAACAATGAGGCTGGGAAATATCTTAAAGACGAAATCCGATATTAAG  
591 ACTGGGTGTAACAATGAGGCTGGGAAATATCTTAAAGACGAAATCCGATATTAAG  
702 CTGATTTGTTGGAACGAGTGAAGTCCAGTCTCTCAGAGAGAAAGCTGTTCCACAC  
531 CTGATTTGTTGGAACGAGTGAAGTCCAGTCTCTCAGAGAGAAAGCTGTTCCACAC  
762 AAGATTGCAAGGATTTGTTTATCCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTG  
471 AAGATTGCAAGGATTTGTTTATCCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTG  
822 GAAGTGTTCATATCAAGTATGAGCAATAGAACTGCAAGCTTCTTGGCTTAA  
411 GAAGTGTTCATATCAAGTATGAGCAATAGAACTGCAAGCTTCTTGGCTTAA  
882 GAAGTGTTCATATCAAGTATGAGCAATAGAACTGCAAGCTTCTTGGCTTAA  
351 GAAGTGTTCATATCAAGTATGAGCAATAGAACTGCAAGCTTCTTGGCTTAA  
942 GAAGTGTTCATATCAAGTATGAGCAATAGAACTGCAAGCTTCTTGGCTTAA  
291 GAAGTGTTCATATCAAGTATGAGCAATAGAACTGCAAGCTTCTTGGCTTAA  
1002 TACCTGTTCTCCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG  
231 TACCTGTTCTCCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG  
1062 CCCTGAATTCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG  
171 CCCTGAATTCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG  
1122 AACTAGATTGTTAGAGTACTACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG  
111 AACTAGATTGTTAGAGTACTACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG  
1182 AAGACTAGTTTCTTCTGATGCCCCCTGAGAGGATTAATTTGTTGTTGTTGTTG  
51 AAGACTAGTTTCTTCTGATGCCCCCTGAGAGGATTAATTTGTTGTTGTTGTTG

RESULT 3  
BI969937/c  
LOCUS  
DEFINITION  
Gm-r1083 Glycine max cDNA clone Gm-r1083-3109.3',  
mRNA sequence.  
ACCESSION  
BI969937  
VERSION  
BI969937.1  
KEYWORDS  
Glycine max (soybean)  
ORGANISM  
Glycine max (soybean)  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

REFERENCE  
1 (bases 1 to 776)  
Vodkin, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Coryell, V.,  
Erpelting, J., Raph, C., Shoop, E., Pardinas, J., Liu, L. and Lewin, H.  
A Functional Genomics Program for Soybean (NSF 9872565)  
Unpublished (1999)  
Contact: Vodkin, L.O., PI, A Functional Genomics Program for  
Soybean (NSF 9872565)  
Lewin, H. A., Director, Keck Center for Comparative and Functional  
Genomics  
University of Illinois  
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA  
Tel: (217) 244-6147  
Fax: (217) 333-4582  
Email: l-vodkin@uiuc.edu  
This clone is available through: Incyte Genomics, 4633 World  
Parkway Circle St. Louis, Missouri 63134. Phone (800) 430-0030 or  
(314) 427-3222 FAX: (314) 427-3324. Web site:  
<http://www.incyte.com/reagents/catalog.jsp?page=clones/collaboratio>  
n/index

## FEATURES

source

Location/Qualifiers  
1..776  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="Gm-r1083-3109"  
/clone="Gm-r1083"  
/note="The library Gm-r1083 is a sequence-driven, reracked set of 4,992 clones selected from cDNA libraries from various tissues and stages of development of soybean. It represents 1117 sequences from the progenitor library Gm-cl009 (from mature roots of 2 month old greenhouse grown 'Williams' soybean plants); 820 sequences from the progenitor library Gm-cl013 (from 2 to 3 week old whole plants of Williams); and 3055 sequences from library Gm-cl028 (from 'Superpod' plants whose seedlings were inoculated with Bradyrhizobium japonicum, courtesy of Dr. Gary Stacey). The 5' ESTs of the source clones from the different progenitor libraries was used to select singletons, or a representative of each contig, which were reracked to form library Gm-r1083. The cDNA clones of the reracked Gm-r1083 library were then sequenced at the 3' end. The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Center for Computational Genomics and Bioinformatics, University of Minnesota, <http://web.abc.umn.edu/biodata/nfsioy/>. Reracking was performed by Incyte Genomics, St. Louis, <http://www.incyte.com>, and 3' sequencing by the Keck Center for Comparative and Functional Genomics, University of Illinois, <http://www.lie.uic.edu/biotech/keck.html>. Note: The corresponding 5' EST from each clone in the Gm-r1083 library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the Incyte Genomics clone ID of the original cDNA library that is also listed under 'OTHER EST'."



Db 539 AGCTGATTGTTGGAACCAAGTGAAGTCCAGTCTCTCAGGAGGAAAGCTGTCAC 480  
 QY 760 ACAAGATTCAAGGATGCTGCTGTTTATCCCTGCTCTTGAAGTCAATCTTCTG 819  
 Db 479 ACAAGATTCAAGGATGCTGCTGTTTATCCCTGCTCTTGAAGTCAATCTTCTG 420  
 QY 820 AFGAAGTTGTTCAATATCAAGTATGAGCAATAGAACTGCAAGCTTCTTGGCTTA 879  
 Db 419 AFGAAGTTATCAATATCAAGTATGAGCAATAGAACTGCAAGCTTCTTGGCTTA 360  
 QY 880 AAGAAGGCTATTGTTGGAATATCTCCGAGCTGCACTGCTGCTTTCAGATTG 939  
 Db 359 AAGAAGGCTATTGTTGGAATATCTCCGAGCTGCACTGCTGCTTTCAGATTG 300  
 QY 940 CAAAAGACCAAGAAATGCGGAGAGCTTATTGTCCTGTTTCCAGCTTCGGGAGA 999  
 Db 299 CAAAAGACCAAGAAATGCGGAGAGCTTATTGTCCTGTTTCCAGCTTCGGGAGA 240  
 QY 1000 GGTACCTGCTCTCCGCTGCTATTTAGTCAAGTGAAGCGGAGCTGAAAGCATGCTTTG 1059  
 Db 239 GGTACCTGCTCTCCGCTGCTATTTAGTCAAGTGAAGCGGAGCTGAAAGCATGCTTTG 180  
 QY 1060 AGCCCTGAATTCCTGTTAAGGCTTCACTACTGAATTTCTGTTACTTGAACAGGCT 1119  
 Db 179 AGCCCTGAATTCCTGTTAAGGCTTCACTACTGAATTTCTGTTACTTGAACAGGCT 120  
 QY 1120 TTAAGTGAATTTGTTAGTACTGTTTGTGACTCTGACTCTAAATAAAGCTTGTCC 1179  
 Db 119 TTAAGTGAATTTGTTAGTACTGTTTGTGACTCTGACTCTAAATAAAGCTTGTCC 60  
 QY 1180 AAAAGACTAGTTTCTTGTGATGCTCCCTGAGCGGATTTTGTGCTGCAACATTAAA 1238  
 Db 59 AAAAGACTAGTTTCTTGTGATGCTCCCTGAGCGGATTTTGTGCTGCAACATTAAA 1

RESULT 5  
 BG648907  
 LOCUS 803 bp mRNA linear EST 24-APR-2001  
 DEFINITION BG648907 HOGA Medicago truncatula cDNA clone PHOGA-24A4 5' end,  
 mRNA sequence.  
 ACCESSION BG648907 GI:13784019  
 VERSION BG648907  
 KEYWORDS  
 SOURCE Medicago truncatula (barrel medic)  
 ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
 Medicago.  
 REFERENCE 1 (bases 1 to 803)  
 AUTHORS Hahn, M.G., Ojaneen-Reubs, T., Samac, D., Town, C.D., Van Aken, S.,  
 Utterback, T., Cho, J. and Fraser, C.M.  
 TITLE ESTs from roots of Medicago truncatula treated with  
 oligogalacturonides of DP 6-20  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Michael G. Hahn  
 Complex Carbohydrate Research Center  
 University of Georgia  
 220 Riverbend Road, Athens, GA 30602-4712, USA  
 Tel: 706-542-4457  
 Fax: 706-542-4412  
 Email: hahn@ccrc.uga.edu  
 G392513e TIGR sequence name: MEMCR02TK More information is  
 available at: www.medicago.org  
 Seq primer: SKmod (CTA GAA CTA GAG GAT CC).  
 FEATURES  
 source location/Qualifiers  
 1. 803  
 /organism="Medicago truncatula"  
 /mol\_type="mRNA"  
 /cultivar="A17"  
 /db\_xref="taxon:3880"  
 /clone="PHOGA-24A4"  
 /tissue\_type="3 day old seedling roots"

/dev stage="24 hours after treatment in the dark at 26 C  
 with 0.5 mg/ml oligogalacturonides (DP 6-20) in the  
 presence of 100 ug/ml Gentamicin"  
 /lab\_host="XLOLR"  
 /clone\_id="HOGA"  
 /note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:  
 XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA  
 was directionally ligated into the Unizap XR vector from  
 Stratagene and packaged using Gigapack III Gold packaging  
 extracts. Plasmids containing cDNA inserts were excised  
 from the recombinant lambda-Zap phage using Ex-assist  
 helper phage and propagated in SOLR cells."

ORIGIN  
 Query Match 47.8%; Score 651.4; DB 12; Length 803;  
 Best Local Similarity 91.3%; Pred. No. 5.3e-109;  
 Matches 724; Conservative 0; Mismatches 66; Indels 3; Gaps 3;

QY 75 TAGTACGAGATATATAGCTGTGAAAGTCCGGAATTGCCAAGATGTTACGGAATTG 134  
 Db 14 TAGTACGAGATATATAGCTGTGAAAGTCCGGAATTGCCAAGATGTTACGGAATTG 72  
 QY 135 ATTGTTAAACCCCATTAATATCTAATATACTTCCGATGTTGTTGCTCCGGGTT 194  
 Db 73 ATTGTTAAACCCCATTAATATCTAATATACTTCCGATGTTGTTGCTCCGGGTT 132  
 QY 195 GCTGCTAACTGAGTGTATGAGAGCCATGCTTATGTTGAAAGACAGAGATGGGTATAGT 254  
 Db 133 GCTGCTAACTGAGTGTATGAGAGCCATGCTTATGTTGAAAGACAGAGATGGGTATAGT 192  
 QY 255 ATGATTGCTGATGAGAGAGAGAGAGGAGGCTTATACACCTGGAAGAGTGTCTCATTTAG 314  
 Db 193 ATGATTGCTGATGAGAGAGAGAGAGGAGGCTTATACACCTGGAAGAGTGTCTCATTTAG 252  
 QY 315 CCAACAAGTGTATATACCTGCAATTTAGCTTCAATGCAAGACAGCCAGGCTTACAAG 374  
 Db 253 CCAACAAGTGTATATACCTGCAATTTAGCTTCAATGCAAGACAGCCAGGCTTACAAG 312  
 QY 375 CTCATTAATTACATGCTGCTCTTATGAGTCTTGAAGAGAAATCATTTCTATTAGCTTTT 434  
 Db 313 CTCATTAATTACATGCTGCTCTTATGAGTCTTGAAGAGAAATCATTTCTATTAGCTTTT 372  
 QY 435 GGAAGTGAAGTGTCTGCAAGATCCTGCTAAGGGAATGAAGGTGCTGTTCAAGAGCT 494  
 Db 373 GGAAGTGAAGTGTCTGCAAGATCCTGCTAAGGGAATGAAGGTGCTGTTCAAGAGCT 432  
 QY 495 GAAGAGATATTGCTAAGAGCCCAATGCTTACATCTTCAACAATTGAAACCTCTGCC 554  
 Db 433 GAAGAGATATTGCTAAGAGCCCAATGCTTACATCTTCAACAATTGAAACCTCTGCC 492  
 QY 555 AATCCCAAGGTTCAATATGAACCACTGCTCAGAGATATGAAGGCTCCGATGGGAAA 614  
 Db 493 AATCCCAAGGTTCAATATGAACCACTGCTCAGAGATATGAAGGCTCCGATGGGAAA 552  
 QY 615 ATTGAATGATTTGTTCTGGGATAGGCACTGCTGTTACATTAACAGGTGCTGAAATAT 674  
 Db 553 ATTGAATGATTTGTTCTGGGATAGGCACTGCTGTTACATTAACAGGTGCTGAAATAT 612  
 QY 675 CTTAAGAGCAGATCCGAATATAAGCTGATTGTTGGAACCAAGTGAAGTCCAGTG 734  
 Db 613 CTTAAGAGCAGATCCGAATATAAGCTGATTGTTGGAACCAAGTGAAGTCCAGTG 672  
 QY 735 CTCTGAGAGGAAGCTGCTCACAAGATTCAGAGATGCTGCTGTTTATCCCT 794  
 Db 673 CTCTGAGAGGAAGCTGCTCACAAGATTCAGAGATGCTGCTGTTTATCCCT 732  
 QY 795 GGTGCTTTGGAAGTCAATCTTGTGATGAAGTGTTCATTAATCAAGTGAAGCAATA 854  
 Db 733 GGTGCTTTGGAAGTCAATCTTGTGATGAAGTGTTCATTAATCAAGTGAAGCAATA 790  
 QY 855 GAAACTGCAAGC 867  
 Db 791 GAAACTGCAAGC 803



RESULT 6  
 AM311624  
 LOCUS  
 DEFINITION 678 bp mRNA linear EST 02-DEC-2001  
 8943b07.y1 Gm-cl025 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
 Gm-cl025-1070 5' similar to TR:O65747 O65747 CYSTEINE SYNTHASE,  
 O-ACETYL-L-SERINE, mRNA sequence.  
 AM311624  
 AM311624.1 GI:6727270  
 EST.  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Glycine max (soybean)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 1 (bases 1 to 678)  
 Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,  
 Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,  
 Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,  
 Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N.,  
 Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,  
 McCann, R., Waterston, R. and Wilson, R.  
 Public Soybean EST Project  
 Unpublished (1999)  
 CONTACT: Shoemaker R./Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: Resgen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntsville, AL 35801 For further information  
 call: (800)-533-4363 or contact via email: ccu@resgen.com  
 Insert Length: 893 Std Error: 0.00  
 High quality sequence stop: 475.  
 Location/Qualifiers  
 1..678  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl025-1070"  
 /tissue\_type="Hypocotyl, 3 day seedlings"  
 /lab\_host="DH10B"  
 /clone\_lib="Gm-cl025"  
 /note="Vector: pT7T3Pac (pT7T3, Pharmacia); Site 1: EcoRI;  
 Site 2: NotI; This cDNA library was constructed from mRNA  
 isolated from hypocotyl tissue of 3 day old seedlings.  
 Complementary DNA was synthesized from mRNA using a poly  
 (dT) primer with a NotI restriction site. EcoRI adapters  
 were ligated to the blunt-ended cDNA fragments followed by  
 digestion with EcoRI and NotI. The cDNA fragments were  
 directionally cloned into the EcoRI-NotI restriction site  
 of the pT7T3-Pac vector. The ligated cDNA fragments were  
 transformed into DH10B host cells (Gibco BRL). This  
 library was constructed by Dr. Randy Shoemaker."

ORIGIN  
 Query Match 47.1%; Score 642; DB 10; Length 678;  
 Best Local Similarity 98.2%; Pred. No. 3e-107;  
 Matches 659; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

QY 331 CTGGCATTTAGTATGCTTCAAGCAGCAGCCAGGGTTACAACTCATATTAATCAATGC 390  
 |||||||  
 DB 8 CTGGCATTTAGTATGCTTCAAGCAGCAGCCAGGGTTACAACTCATATTAATCAATGC 67  
 |||||||  
 QY 391 CTGCTTTATAGAGTCTTGAGAGAGAAATCATTTATAGCTTTTGAGAGCTGAGTTGCTC 450  
 |||||||  
 DB 68 CTGCTTTATAGAGTCTTGAGAGAGAAATCATTTATAGCTTTTGAGAGCTGAGTTGCTC 127  
 |||||||  
 QY 451 TGACAGATCTCTGTAAGGAATGAAGTGTCTTTCAGAAAGCTGAAGATATATGGCTA 510

DB 128 TGACAGATCTCTGTAAGGAATGAAGTGTCTTTCAGAAAGCTGAAGATATATGGCTA 187  
 |||||||  
 QY 511 AGACGCCCAATGCTTACATTAATCTTCAACAATTTGAACAAACCTGCCAATCCCAAGTTCAAT 570  
 |||||||  
 DB 188 AGACGCCCAATGCTTACATTAATCTTCAACAATTTGAACAAACCTGCCAATCCCAAGTTCAAT 247  
 |||||||  
 QY 571 ATGAACCACTGCTCCAGAGATATGGAAGGCTCCGATGGGAAATATGATGCAATTTGTTT 630  
 |||||||  
 DB 248 ATGAACCACTGCTCCAGAGATATGGAAGGCTCCGATGGGAAATATGATGCAATTTGTTT 307  
 |||||||  
 QY 631 CTGGATAGGCACTGCTGTATCAATATACAGGCTCTGGAATAATATCTTAAAGAGCAATTC 690  
 |||||||  
 DB 308 CTGGATAGGCACTGCTGTATCAATATACAGGCTCTGGAATAATATCTTAAAGAGCAATTC 367  
 |||||||  
 QY 691 CGAATATAAGCTGATTTGCTGTGGAACCAAGTTGAAGTCCAGTCTCTCAGAGAAAGC 750  
 |||||||  
 DB 368 CGAATATAAGCTGATTTGCTGTGGAACCAAGTTGAAGTCCAGTCTCTCAGAGAAAGC 427  
 |||||||  
 QY 751 CTGGTCCACACAGATTCAGAGGATTTGCTGCTTTTATCCCTGCTGTGGAAGTCA 810  
 |||||||  
 DB 428 CTGGTCCACACAGATTCAGAGGATTTGCTGCTTTTATCCCTGCTGTGGAAGTCA 487  
 |||||||  
 QY 811 ATCTTCTGATGAAGTTGTTCAATATCAAGTATGAGCAATAGAACTGCAAGCTTC 870  
 |||||||  
 DB 488 ATCTTCTGATGAAGTTGTTCAATATCAAGTATGAGCAATAGAACTGCAAGCTTC 547  
 |||||||  
 QY 871 TTGGCTTAAGAAGGCTTATTTGGAATATCTCCGAGCTGACGCTGCTGCTT 930  
 |||||||  
 DB 548 TTGGCTTAAGAAGGCTTATTTGGAATATCTCCGAGCTGACGCTGCTGCTT 607  
 |||||||  
 QY 931 TTCAATTTCAAAAAGACCAAAAT-GCCGGAAGCTTATTTGCTGCTTTTCCAGC 989  
 |||||||  
 DB 608 TTCAATTTCAAAAAGACCAAAATGCGCGGAAGCTTATTTGCTGCTTTCCAGC 667  
 |||||||  
 QY 990 TTCCGGGAGAG 1000  
 |||||||  
 DB 668 TTCCGGGAGAG 678

RESULT 7  
 CF513886  
 LOCUS  
 DEFINITION 929 bp mRNA linear EST 09-SEP-2003  
 CABud0007\_IIIF H10 Vitis vinifera cv. cabernet sauvignon (Clone 8)  
 Bud - CABUD Vitis vinifera cDNA clone CABud0007\_IIIF\_H10 5', mRNA  
 sequence.  
 CF513886  
 CF513886.1 GI:34545654  
 EST.  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Vitis vinifera  
 Vitis vinifera  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; Vitaceae; Vitis.  
 1 (bases 1 to 929)  
 Goes da Silva, F., Iandolo, A., Lim, H., Baek, J., Jones, K. and  
 Cook, D.  
 Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'  
 berries at various developmental stages  
 Unpublished (2003)  
 CONTACT: Douglas Cook, PhD  
 CARS Genome Facility  
 UC Davis, Plant Pathology  
 One Shields Ave, Davis, CA 95616, USA  
 Tel: 530 754 6561  
 Fax: 530 754 6617  
 Email: drcook@ucdavis.edu  
 Seq primer: ACGGTACCGGACATATGCC.  
 Location/Qualifiers  
 1..929  
 /organism="Vitis vinifera"  
 /mol\_type="mRNA"  
 /cultivar="Cabernet Sauvignon (Clone 8)"

FEATURES  
 source

/db xref="taxon:29760"  
/clone="Cabud007\_11IF\_H10"  
/sex="Hermaphrodite"  
/dev stage="Pre-bloom (10-11 days before bloom)"  
/lab\_host="DH5alpha"  
/clone\_lib="Vitis vinifera cv. cabernet sauvignon (Clone 8) Bud - CABUD"  
/note="Organ: Bud; Vector: pDNR; Site 1: SfiI; Site 2: SfiI; CABUD is a cDNA library of Vitis vinifera cv. Cabernet Sauvignon Clone 8 dissected buds. Samples were collected May 13, 2002 from pre-bloom plants (10-11 days before bloom), pre-veraison. Sampled vines were located at the University of California, Davis, Experimental Vineyard. cDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows:  
5'-AAGCAGTGTATCAACGACAGAGTGGCCATTACGGCCGG-3' and 5'-ATTCTAGAGCGCGGCGGCGACATG-dT(30)NM-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

ORIGIN

Query Match 46.7%; Score 636.4; DB 14; Length 929;  
Best Local Similarity 80.8%; Pred. No. 2.7e-106;  
Matches 742; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

QY 150 TTATATATCTTAATAAATCTTGGAGTGTGTGTCCTGCTGCTTAACCTGAG 209  
DB 1 TTATATATCTGAACAATGTGTAGATGGCTGTGTAGCTCTATTTGCTGCAAGCTTGA 60  
QY 210 TTGATGAGCCATGCTCTAGTGTGAAGACAGAGTGGTATAGTATGATGCTGATGCA 269  
DB 61 ATGATGAGCCTTGTCTAGTGTGAAGACAGAGTGGTATAGTATGATGATGATGCA 120  
QY 270 GAAGAGAGGAGCTTATACACCTGGAAGAGTGTCTCTATGAGCCCAACAGTGTAT 329  
DB 121 GAAGAGAGGAGCTTATACACCTGGAAGAGTGTCTCTATGAGCCCAACAGTGTAT 180  
QY 330 ACTGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 389  
DB 181 ACTGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240  
QY 390 CCGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 449  
DB 241 CCGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300  
QY 450 CTGACAGATCTCTGTAAGGATGAAGAGTGTCTGTTCAAGAGCTGAAGATATGCT 509  
DB 301 CTGACAGATCTCTGTAAGGATGAAGAGTGTCTGTTCAAGAGCTGAAGATATGCT 360  
QY 510 AAGAGCCCAATGCTCTACATCTTCAACATTTGAAAACCTGCCAATCCCAAGTTCAT 569  
DB 361 AAGAGCCCAATGCTCTACATCTTCAACATTTGAAAACCTGCCAATCCCAAGTTCAT 420  
QY 570 TATGAACAACCACTGTGTCAGAGATATGGAAGGCTCCGATGGAAAATGATGATGCTT 629  
DB 421 TATGAACAACCACTGTGTCAGAGATATGGAAGGCTCCGATGGAAAATGATGATGCTT 480  
QY 630 TCTGGAGATAGGCACTGTGTGTACATTAACAGGTGTGTGAAAATATCTTAAGAGCAAT 689  
DB 481 TCTGGAGATAGGCACTGTGTGTACATTAACAGGTGTGTGAAAATATCTTAAGAGCAAT 540  
QY 690 CCGAATATTAAGCTGATGTTGTGTGGAACCAAGTGAAGATCCAGTCTCTCAGAGGAAAG 749  
DB 541 TCTGAACATTAAGCTGATGTTGTGTGGAACCAAGTGAAGATCCAGTCTCTCAGAGGAAA 600  
QY 750 CCGTGTCCACACAGATTAAGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 809  
DB 601 CCGTGTCCACACAGATTAAGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660  
QY 810 AATCTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 869  
DB 661 AATCTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720

QY 870 CTTCGGCTTAAGAGCCCTATTTGTGGAAATCTTCCGAGCTGCAGCTGCTGCT 929  
DB 721 CTTCGATTAAGAGCCCTATTTGTGGAAATCTTCCGAGCTGCAGCTGCTGCT 780  
QY 930 TTTCAGATTGCAAAAAGACCAAGAAATATGCTGGGAAAGCTTATTTGCTGCTTCC 989  
DB 781 ATTAAGTTAGCAAAAGAGCCCAAGAAATATGCTGGGAAAGCTTATTTGCTGCTTCC 840  
QY 990 TTGGGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1049  
DB 841 TTGGGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
QY 1050 ATGACTTTGAGCCCTGA 1067  
DB 901 ATGCTGTTTGAACCTGA 918

RESULT 8  
CA801673  
LOCUS  
DEFINITION  
CA801673 632 bp mRNA linear EST 05-DEC-2002  
sac16d04.y1 Gm-c1036 Glycine max cDNA clone SOYBEAN CLONE ID:  
Gm-c1036-13975 5' similar to TR:023733 023733 O-ACETYL SERINE (THIOL)  
LYASE.; mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
EST.  
Glycine max (soybean)  
Glycine max  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

REFERENCE  
AUTHORS  
1 (bases 1 to 632)  
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,  
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,  
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,  
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,  
Schurk, R., Ritter, E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M.,  
McCann, R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R./Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
This clone is available through: ResGen, Invitrogen Corp. 2130  
South Memorial Parkway Huntville, AL 35801 For further information  
call: (800)-533-4363 or contact: ccu@resgen.com web site:  
www.resgen.com  
Seq primer: -40RP from Gibco  
High quality sequence stop: 448.

FEATURES  
source  
1. 632  
location/Qualifiers  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="SOYBEAN CLONE ID: Gm-c1036-13975"  
/tissue\_type="somatic embryos cultured on MSD 20"  
/lab\_host="DH10B"  
/clone\_lib="Gm-c1036"  
/note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; This  
cDNA library was constructed from mRNA isolated from  
somatic embryos (age ranging from 2 months to 9 months)  
cultured on MSD 20. The library was prepared using the  
Life Technologies pSuperScript cDNA library construction  
kit. Complementary DNA was synthesized from mRNA using a  
poly (dT) sequence with a NotI restriction site. SalI  
linkers adapters were ligated to the blunt-ended cDNA  
fragments followed by NotI digestion. The cDNA fragments

were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. This library was constructed in the laboratory of Dr. Lila Vodkin by Anu Khanna at the University of Illinois at Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

# ORIGIN

Query Match 46.3%; Score 630.4; DB 14; Length 632;  
Best Local Similarity 99.8%; Pred. No. 4e-105;  
Matches 631; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY	13	GTAGATAGCCGATGCTTGTCTAGTGTGTCAGTCACTTCTCTCTCAAGTCAAGCTT	72
Db	1	GTAGATAGCCGATGCTTGTCTAGTGTGTCAGTCACTTCTCTCTCAAGTCAAGCTT	60
OY	73	TGTAGTACGATATATATGCTGTGTAAGGTCGGAATTCGCAAGATGTTACGGAAT	132
Db	61	TGTAGTACGATATATATGCTGTGTAAGGTCGGAATTCGCAAGATGTTACGGAAT	120
OY	133	TGATGTGTAACCCCAATAGTATATCTAATAAATCTGCGATGTGTGTGCGCGG	192
Db	121	TGATGTGTAACCCCAATAGTATATCTAATAAATCTGCGATGTGTGTGCGCGG	180
OY	193	TTCCTGCTAACTGAGTGTGAGGAGCCATGCTCTAGTGTGTAAGGACAGATGGGTATA	252
Db	181	TTCCTGCTAACTGAGTGTGAGGAGCCATGCTCTAGTGTGTAAGGACAGATGGGTATA	240
OY	253	GTATGATGTCTGATGACAGAGAGAGGACTTATCACTGGAAGAGTGTCTCATTTG	312
Db	241	GTATGATGTCTGATGACAGAGAGAGGACTTATCACTGGAAGAGTGTCTCATTTG	300
OY	313	AGCCAACAAGTGTATATCTGCAATGATTAAGCTTCATGCGACAGCGGAGTTACA	372
Db	301	AGCCAACAAGTGTATATCTGCAATGATTAAGCTTCATGCGACAGCGGAGTTACA	360
OY	373	AGCTCATATTAATCAATGCTCTCTTCTATGAGTCTTGAAGAGATCATTTCTAAGCTT	432
Db	361	AGCTCATATTAATCAATGCTCTCTTCTATGAGTCTTGAAGAGATCATTTCTAAGCTT	420
OY	433	TTCGAGCTGAGTGTGCTGACAGATCTCTGCTAAGGGAATGAAGTGTCTGCAAGG	492
Db	421	TTCGAGCTGAGTGTGCTGACAGATCTCTGCTAAGGGAATGAAGTGTCTGCAAGG	480
OY	493	CTGAAGAGATATTGCTTAAGAGCCCAATGCTCTACATCTTCAACAATTTGAACCTG	552
Db	481	CTGAAGAGATATTGCTTAAGAGCCCAATGCTCTACATCTTCAACAATTTGAACCTG	540
OY	553	CCATATCCCAAGGTTCAATTAAGAACCACTGCTCCAGAGATATGAAAGGCTCCGATGGA	612
Db	541	CCATATCCCAAGGTTCAATTAAGAACCACTGCTCCAGAGATATGAAAGGCTCCGATGGA	600
OY	613	AAATGATGATTTGTTCTTCTGAGATAGGCACT	644
Db	601	AAATGATGATTTGTTCTTCTGAGATAGGCACT	632

RESULT 9  
CB974516 910 bp mRNA linear EST 01-MAY-2003  
LOCUS CAB30005.111b\_Fb\_F10 Cabernet Sauvignon Berry Stage I - CAB3 Vitis  
DEFINITION Vitis vinifera cDNA clone CAB30005.111b\_Fb\_F10 5', mRNA sequence.  
ACCESSION CB974516  
VERSION CB974516.1 GI:30297722  
KEYWORDS EST  
SOURCE Vitis vinifera  
ORGANISM Vitis vinifera  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; Vitaceae; Vitis.  
REFERENCE 1 (bases 1 to 910)  
Goes da Silva, F., Iandolo, A., Lim, H., Baek, J., Jones, K. and Cook, D.

TITLE  
JOURNAL  
COMMENT  
Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon' berries at various developmental stages  
Unpublished (2003)  
Contact: Douglas Cook, PhD  
CAES Genome Facility  
UC Davis, Plant Pathology  
One Shields Ave, Davis, CA 95616, USA  
Tel: 530 754 6561  
Fax: 530 754 6617  
Email: drcoc@ucdavis.edu  
Seq primer: ACGGTACCGACATATGCC.  
Location/Qualifiers

## FEATURES

source

1. 910  
/organism="Vitis vinifera"  
/mol\_type="mRNA"  
/cultiivar="Cabernet Sauvignon"  
/db\_xref="taxon:29760"  
/clone="CAB30005.111b\_Fb\_F10"  
/sex="Hermaphrodite"  
/dev\_stage="Berry stage 1"  
/lab\_host="DH5alpha"  
/clone\_11b="Cabernet Sauvignon Berry Stage I - CAB3"  
/note="Organ: Berry; Vector: pDNR; Site 1: Sfil; Site 2: Sfil; CAB3 is a cDNA library of Vitis vinifera 'Cabernet Sauvignon' Clone 8 berries. Samples were collected after berry set from field-grown vines during stage 1 of berry growth, 17 days after full bloom. The average berry size was 6 millimeters. Sampled vines were located at the University of California, Davis, Experimental Vineyard. cDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows: 5'-AAGCAATGATCAAGCAGAGTGGCATTAAGGCGG-3' and 5'-ATTCTAGAGCGCGAGCGCGGACATG-DT(30)NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

## ORIGIN

Query Match 46.2%; Score 629.2; DB 14; Length 910;  
Best Local Similarity 81.7%; Pred. No. 5.6e-105;  
Matches 738; Conservative 0; Mismatches 164; Indels 1; Gaps 1;

OY	111	ATTGCCAAGATGTTACGGAATGATGTAATAAACCCCATTAATATCTAATAAATCTT	170
Db	8	ATTGCCAAGATGTTACGGAATGATGTAATAAACCCCATTAATATCTAATAAATCTT	67
OY	171	GCGGATGTTGTGTGTCGCGGCTGCTGCTAACTGAGTGTATGAGCCATGCTCTAGT	230
Db	68	GCGGATGTTGTGTGTCGCGGCTGCTGCTAACTGAGTGTATGAGCCATGCTCTAGT	127
OY	231	GTGAAGGACAGAGTGGGTATAGTATGATGCTGATGACAGAGAGAGGAGCTTACACA	290
Db	128	GTGAAGGACAGAGTGGGTATAGTATGATGCTGATGACAGAGAGAGGAGCTTACACA	187
OY	291	CCTGGAAGAGTGTCTCTCAATGAGCCCAAGTGTATATCTGSCATTTAGAGCTTTC	350
Db	188	CCTGGAAGAGTGTCTCTCAATGAGCCCAAGTGTATATCTGSCATTTAGAGCTTTC	247
OY	351	ATGGCAGCAGCCAGGCTTACAGCTCATTAATTAACAATGCTGCTTCTATGAGCTT	410
Db	248	ATGGCAGCAGCTAAGGCTTATAGCTCATTAACAATGCTGCTTCTGATGAGCTT	307
OY	411	AGAAGATCATTTATAGCTTTTGGAGCTGAGTGTCTTGAAGATCTCTGTAAGGA	470
Db	308	AGAAGATCATTTCTCGAGCTTTTGGGCTGAGTGTCTTCAAGATCCAGCCAGGGC	367
OY	471	ATGAAGGCTGCTTCAAGAGGCTGAAGAGATATTGCTTAAGAGCCCAATGCTTACATA	530
Db	368	ATGAAGGCTGCTTCAAGAGGCTGAAGAGATATTGCTTAAGAGCCCAATGCTTACATA	427
OY	531	CTTCAACAATTTGAACCCCTGCAATCCCAAGTTCATTAAGAACCACTGCTCCAGAG	590
Db	428	CTTCAACAATTTGAACCCCTGCAATCCCAAGTTCATTAAGAACCACTGCTCCAGAG	487



Dy	591	ATATGGAAGGCCTCCGATGGCAAAATTGATGTCATTTGTTCGCGATAGGCACCTGTGCT	650
Dy	488	ATCTGGAGAGGCCACTGTGGGAAAAGTGTGATGCTTTTGCTCTGGCAATAGGCACTGGAGCT	547
OY	651	ACAATAACAGGTGCTGCGAAAAATATCTTAAGAGCAGAATCCGAATATAAGCTGATTGCT	710
Dy	548	AACATPACAGGTGCAGGGAAGTTCCTCAAGAGCCMAAATTCGACCTTAAGCTGTATGCT	607
OY	711	GTEGAACCAGTTGAAAGTCCAGTGTCTCAGAGGAAAGCCTGTCCACACAGATTGCA	770
Dy	608	GTAGAACCTGTGMAAGTGCAGTACTGTCTGGAGMAAACCTGGCCCGCATPAAGATCCA	667
OY	771	GCGATTGTGCTGTGTTTATCCCTGTGTCTTGGAAGTCAACTCTTGATGAAGTTGTT	830
Dy	668	GCAATTGGCGCTGTTCATCCTCGAAGTTTAGATGTCAATTTGCTTGATGAAGTTGTT	727
OY	831	CAATATCAAGTGATGAAGCAATAGAACTGCCAAAGCTTCTGCGCTTAAGAGGCGCTA	890
Dy	728	CAAGTTCAAGTGAGAAGCTGTGAAACTGCCAAGCTTCTTGCAATTGAAGAGGTTTG	787
OY	891	TTTGTGGGAATATCTCCGAGCTGCAGCTGCTGTGCTTTCAGATTGCCAAAAAGACCA	950
Dy	788	CTGCTTGGGATATATCATCTGTGTGTGCTGCAGCTGTGCAATAAAGTAGCAAGAGGCCA	847
OY	951	GAAATGCCCGGAAGCTTATT-GTTGCCGTTTTTCCAGGCTCGGGAGAGGTAACCTGTC	1009
Dy	848	GAAATGCTGGGAACCTCATTTGGATCMGTTTTCCCAAGCTTGGAGAGCGTTATCTGTC	907
OY	1010	CTC 1012	
Dy	908	CTC 910	

```

RESULT 10
LOCUS      BF003458
DEFINITION ESTf31956 KVI Medicago truncatula cDNA clone pKV1-5N8, mRNA
ACCESSION  BF003458
VERSION    BF003458
KEYWORDS   GI:10703733
SOURCE     Medicago truncatula (barrel medic)
ORGANISM   Medicago truncatula
            Burkarya; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE  1 (bases 1 to 724)
AUTHORS   VandenBosch,K., Endre,G., Hur,J., Moore,J., Beremand,P., Ellis,L.,
            Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S., Holt,I.E. and
            Fraser,C.M.
TITLE     ESTs from roots of Medicago truncatula 24 hours after inoculation
            with Sinorhizobium meliloti
JOURNAL    Unpublished (1999)
COMMENT    Contact: VandenBosch K
            Department of Plant Biology
            University of Minnesota
            220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
            Tel: 612 624 2755
            Fax: 612 625 1738
            Email: kvandenb@chs.umn.edu
            Texas A&M University name: T267268e TIGR sequence name: MT1AP76TK
            More information is available at: http://chryslie.tamu.edu/medicago
            Seq primer: SKmod (CTA gAA CTA gtg gAT CC).

FEATURES
    source
        1..724
            /organism="Medicago truncatula"
            /mol_type="mRNA"
            /cultivar="genotype A17"
            /db_xref="taxon:3880"
            /clone="pKV1-5N8"
            /tissue_type="Seedling roots"
            /dev_stage="24 hours post-inoculation with Sinorhizobium

```

```

mel10ct1"
/lab_host="E. coli strain X10LR"
/clone_lib="KV1"
/notes="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the Unizap XR vector from
Stratagene and packaged using Gigapack II Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-Zap phage using Ex-assist
helper phage and propagated in X10LR cells."

```

Query Match	46.0%;	Score 627.2;	DB 10;	Length 724;
Best Local Similarity	91.9%;	Pred. No. 1.4e-104;		
Matches	662;	Conservative	0;	Mismatches 58; Indels 0; Gaps 0;
QY	96	GTTGAAGGTCGGGAATTGCCAAGAATGCTTACGGAATTGATTTGGTAAACCCCATTAAGTA	155	
Db	5	GGTCGAAGGTCGGAATTGCTTAAAGATGTTACAGAAATTGATTTGGCAAAACCCCGTTAGTG	64	
QY	156	TATCTAATAATAACTTGGCGATGGTTGTGTGCCCCGGGTGCTGCTAAACTGGAGTTGATG	215	
Db	65	TATCTAATAATAGACTTGGCAGATGGTTGTGTGCTGCTGCACTTGTCTGCCAACTGGAATTGATG	124	
QY	216	GAGCCATGCTCTAGTGTGAAGGACAGGATTTGGGTATAGTATGATTTGCTGATGACAGAGAG	275	
Db	125	GAGCCATGCTCTAGTGTGAAGGACAGGATTTGGGTATAGTATGATTTGCTGATGACAGAGAG	184	
QY	276	AAGGACTTATACACACCTGGAAAAGATGTCTCTCATTTGAGCCCAAGAGTGTATATCTGGC	335	
Db	185	AAGGGCTTATACACCTGGACAGAGTGTCTCTCATTTGAAACCAAGAGTGTATATCTGGC	244	
QY	336	ATTGATTAGCCTTCATGCGACAGCAGCCAGGGGTACAGCTCATTAATTACATGCTGCT	395	
Db	245	ATTGATTAGCCTTCATGCGACAGCAGCCAGGGGTACAGCTCATTAATTACATGCTGCT	304	
QY	396	TCTATAGTCTTGAGAGAAAGATCATTTCTATTAGCTTTTGGAGCTGAGTTGGTTCTGACA	455	
Db	305	TCAATGAGTCTTGAAAGAAAGATCATTTCTATTAGCTTTTGGAGCTGAGTTGGTTTGAACA	364	
QY	456	GATCCTGCTAAGGGAATGAAGGTGCTGTTCAGAAAGCTGAAGAGATATTTGCGTAAGACG	515	
Db	365	GACCTGCAAGGGAATGAAGGAGCTGTTCAAGAACTGAAGAGCTGTTGCGTAAGACA	424	
QY	516	CCCAATGCTACATACTTCAACAATTTGAAAAACCTGCCAATCCCAAGGTTCAATTATGAA	575	
Db	425	CCCAATGCTACATACTTCAACAATTTGAAAAACCTGCCAATCCCAAGGTTCAATTATGAA	484	
QY	576	ACCACTGTGTCAGAGATATGGAAGGCTCCGATGGGAAAAATTGATGCATTTTGTCTGGG	635	
Db	485	ACCACCGGTCCAGAGATATGGAAGGCAACGATGGGAAAAATTGACGCATTTTGTCTGGG	544	
QY	636	ATAGGCACTGTGTGTACATTAACAGGTGCTGGAAAAATATCTTAAAGACAGATCCGAAT	695	
Db	545	ATAGGTACTGTGTGTACATTAACCGGTGCTGGAAAAATATCTTAAAGACCAAAATTTCCAA	604	
QY	696	ATAAGCTGATTGGTGTGGAACCAAGTTGAAGTCCAGTGTCTCAGGAGGAAAGCCTGGT	755	
Db	605	ATAAGCTGATTGGTGTGGAACCAAGTGAAGTCCAGTGTCTCAGGAGGAAAGCCTGGT	664	
QY	756	CCACACAAGATTCAAGGGAATTGGTGTCTGTTTATCCCTGGTGTCTTTGGAAGTCAATCTT	815	
Db	665	CCTCAACAAGATTCAAGGGAATTGGTGTCTGCTTGTCTTGTCTGTCTTTAGAAAGTTAATCTT	724	

SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Solanum tuberosum (potato)  
Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum.  
1 (bases 1 to 791)  
van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Chiemingo, A.,  
Bougri, O., Buell, C.R., Romming, C., Tanksley, S. and Baker, B.  
Generations of ESTs from sprouting potato eyes  
Unpublished (2000)  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
This clone can be obtained from the University of Arizona Genomics  
Institute. Orders can be made through URL:  
http://genome.arizona.edu/orders/  
Seq primer: M13F-R.  
Location/Qualifiers

FEATURES  
source  
1..791  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="cST311N11"  
/tissue\_type="sprouting eyes from tubers"  
/dev\_stage="12-14 weeks post harvest"  
/lab\_host="SOLR"  
/clone\_lib="cSTS"  
/note="Vector: Bluescript SK(-); Site\_1: EcoRI; Site\_2:  
XhoI; Various sizes of sprouting eyes (2mm to 15mm) were  
taken from tubers. The tubers were incubated at 26C in the  
dark for 2-3 weeks prior to sprouting. The eyes were  
frozen in liquid nitrogen immediately upon removal from  
tubers."

ORIGIN

Query Match 46.0%; Score 626.2; DB 12; Length 791;  
Best Local Similarity 89.7%; Pred. No. 2.1e-104;  
Matches 695; Conservative 0; Mismatches 78; Indels 2; Gaps 2;

51 TCCCTGTTCTCAAGTCAAGCTTGTAGTGAAGATATATAGCTGTGAAGGCTCGGA 110  
12 TCCATTTCTCAACACACACACAGTAGTAGAGCTGATACATGCGCTCGAAA-GTCTGA 70  
111 ATTGCCAAGATGTACGGAATGATGTGTAACCCCATAGTATATCTAATAACTT 170  
71 ATTGCTAAGATGTACAGAAATGATGCAAAACCCGTTAGTGTATCTAATAGACTT 130  
171 GCGGATGTTGTTGTTGCCCCGGTGTCTTAACCTGAGTGTAGGACCATGCTTAGT 230  
131 GCAGATGTTGTTGTTGCTCGAGTGTCTGCCAAACTGGAATGATGAGCCATGCTTAGT 190  
231 GTGAAGCAGAGATTGGGTATAGTATGATTCCTGATGACAGAGAGAGGACTTATACA 290  
191 GTAAAGACAGAGATTGGCTACAGTATGATTCCTGATGACAGAGAGAGGCTTATACA 250  
291 CCTGGAAGAGTGTCTCTCATTTAGGACCAACAAGTGTATATCTGCAATTGATAGCTTC 350  
251 CCTGGAAGAGTGTCTCTCATTTAGGACCAACAAGTGTATATCTGCAATTGATAGCTTC 310  
351 ATGGCAGAGCCAGGGGTACAGCTCATTAATCAATGCTTCTATAGTCTTAG 410  
311 ATGGCAGAGCCAGGGGTACAGCTCATTAATCAATGCTTCTATAGTCTTAG 370  
411 AGAAGATCATTTATTTAGCTTTTGAAGTGAAGTGTCTTCTGACAGATCTGTAAGGA 470  
371 AGAAGATCATTTATTTAGCTTTTGAAGTGAAGTGTCTTCTGACAGATCTGTAAGGA 430  
471 ATGAAGTGTCTTCTGAGAGGCTGAAGATATTTGGCTAAGAGCCCAATGCTTACATA 530  
431 ATGAAGTGTCTTCTGAGAGGCTGAAGATATTTGGCTAAGAGCCCAATGCTTACATA 490

531 CTTCAACAATTGAAAACCTGCCAATCCCAAGTTCAATTATGAAACCACTGTTCCAG 590  
491 CTTCAACAATTGAAAACCTGCCAATCCCAAGTTCAATTATGAAACCACTGTTCCAG 550  
591 ATATGAAAAGGCTCCGATGGAATAATGATGATTTGTTCTGGATAGGCACTGTGT 650  
551 ATATGAAAAGGCAAGGATGGAATAATGATGATTTGTTCTGGATAGGCACTGTGT 610  
651 ACAATACAGCTGCTGGAATAATTTTAAGAGCAGAAATCCGATATAAAGCTGA-TTGG 709  
611 ACCATACCGGTGCTGGAATAATTTTAAGAGCAGAAATCCGATATAAAGCTGAATTGG 670  
710 TGTGAACCAAGTTGAAGTCCAGTCTCTCAGAGAGAAAGCTGTCACACAGATTCA 769  
671 TGTGAACCAAGTTGAAGTCCAGTCTCTCAGAGAGAAAGCTGTCACACAGATTCA 730  
770 AGGATGTTGCTGTTTATCCCTGCTGTTGGAAGTCAATCTTGTATGAA 824  
731 AGGATGTTGCTGTTTATCCCTGCTGTTGGAAGTCAATCTTGTATGAA 785

RESULT 12  
CK260011 994 bp mRNA linear EST 12-DEC-2003  
LOCUS  
DEFINITION  
EST706089 potato abiotic stress cDNA library Solanum tuberosum cDNA  
clone POAB228 5' end, mRNA sequence.  
ACCESSION  
CK260011  
VERSION  
CK260011.1 GI:39816989  
KEYWORDS  
EST.  
SOURCE  
Solanum tuberosum (potato)  
ORGANISM  
Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum.  
1 (bases 1 to 994)  
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.  
Generation of ESTs from abiotic stressed potato tissue  
Unpublished (2003)  
Other ESTs: EST706088 EST706090 EST706091  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from TIGR via potato@tigr.org  
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES  
source  
1..994  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="POAB228"  
/tissue\_type="abiotic stress treated leaf and root tissue"  
/lab\_host="DHL08-Tona"  
/clone\_lib="potato abiotic stress cDNA library"  
/note="Vector: PCMVSPORT6.1; Site\_1: EcoRI; Site\_2: NotI;  
supplier: Solanum tuberosum var. Kennebec plants were  
grown from cuttings on a 16hr light/8 hr dark cycle at 25  
C for 3-4 weeks. Abiotic stress conditions were applied to  
four separate sets of plants. Set 1 involved saturation of  
the soil with 150 mM NaCl and tissues were harvested at  
following application of the salt stress (leaves: 2hr,  
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).  
Set 2 were grown under the standard conditions and then  
were water stressed by withdrawal of further watering  
applications. Drought stressed plants were harvested after  
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d  
and 5d). Set 3 were grown under the standard conditions  
and then were cold stressed by placement at 4 C. Cold  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d. Set 4 were grown under the standard conditions and  
then were heat stressed by placement at 35 C. Heat

stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

ORIGIN

Query Match 45.9%; Score 625.4; DB 14; Length 994;  
Best Local Similarity 77.6%; Pred. No. 2.6e-104;  
Matches 769; Conservative 0; Mismatches 221; Indels 1; Gaps 1;

67 AAGCTTTAGTGAAGAGATATATGCTGTTGAAGGTCGGAATGCCAAGATGTTA 126  
5 AATCCTAGATTAACCTGATATATGCGGGGGAAGATGGAATGCCAAGATGTTA 64  
127 CGGAATGATGTTAAACCCATAGTATATCTAAATAACTTGGGATGTTGTTG 186  
65 CTGAATGATGTTAACTCTCTTGTATACCTGATATATGTTGATGGGTGTTG 124  
187 CCGGGGTGCTGCTAACTGAGTGTGAGGCCATGCTCTAGTGTGAAGACAGATTG 246  
125 CACGTGTTGCTGCCAAGCTGGAAGCATGAGCCATGCTCTAGTGTGAAGATGAGTCG 184  
247 GGTATAGTATGATGCTGATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 306  
185 GTTATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 244  
307 TCATTGAGCCCAAG 366  
245 TCATTGAGCCCAAG 304  
367 GTTACCAAGCTCATATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 426  
305 GCTACCAAGCTCATATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 364  
427 TAGCTTTGAGCTGAGTGTGTTGCTGACAGATCTGTTAGAGAGAGAGAGAGAGAG 486  
365 GTGCTTTCGCTGCTGAGTGTGTTGCTGACAGATCTGTTAGAGAGAGAGAGAGAG 424  
487 AGAAGCTGAAGAGATATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 546  
425 CAAAGGCTGAAGAGATATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 484  
547 ACCCTGCCAATCCCAAGGTTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 606  
485 ACCCTGCCAATCCCAAGGTTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 544  
607 ATGGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 666  
545 ATGGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 604  
667 GAAATATCTTAAG 726  
605 GCAAGTATTTGAAG 664  
727 GTCCAGTCTCTCAAG 786  
665 GTGCTATCTTCTGCTGAG 724  
787 TTATCCCTGCTGCTGAG 846  
725 TCATTCCTGCTGCTGAG 784  
847 AAGCAATAG 906  
785 AATCCATAG 844  
907 CCGGAGCTGAG 966  
845 CTGGTCTGCTGCTGAG 904  
967 TTATGTTGCTGCTGAG 1026

Db 905 TCATTGTG-TTATTTCCAGCTTCGAGAGAGAGATATCTTCTCTACTCTCGAAA 963  
QY 1027 CAGTGAAGCGGAG 1057  
Db 964 CTGTCAAG 994

RESULT 13  
BE660317 653 bp mRNA linear EST 06-SRP-2000  
LOCUS BE660317  
DEFINITION 2-D7 GmaxSC Glycine max cDNA, mRNA sequence.  
ACCESSION BE660317  
VERSION BE660317.1 GI:9986209  
KEYWORDS EST.  
SOURCE Glycine max (soybean)  
ORGANISM Glycine max

REFERENCE 1 (bases 1 to 653)  
AUTHORS Harris, N., Chapman, B.P. and Gijzen, M.  
TITLE Gene expression in developing soybean seed coats  
JOURNAL Unpublished (2000)  
COMMENT Contact: Gijzen M  
Agriculture and Agri-Food Canada  
1391 Sandford Street, London, Ontario, Canada N5V 4T3  
Tel: 519 457 1470  
Fax: 519 457 3997  
Email: gijzenm@agr.ca.

FEATURES

source 1.653  
/organism="Glycine max"  
/mol\_type="mRNA"  
/cultivar="Harosoy 63"  
/db\_xref="taxon:3847"  
/tissue\_type="Seed coats"  
/lab\_host="E. coli strain XL0R"  
/clone\_lib="GmaxSC"  
/note="Vector: pBK-CMV; Site 1: EcoRI; Site 2: XhoI; This cDNA library was constructed from polyA+ enriched mRNA from green seed coats in mid to late developmental stage, average fresh weight 250 mg per seed. Traces of pod and embryo tissue also present. Complementary DNA was synthesized from mRNA using an XhoI-poly(dT) linker-primer. EcoRI adapters were ligated to the blunt-ended cDNA fragments and the products were digested with XhoI for directional cloning into lambda ZAP Express vector. This lambda library was amplified once using E. coli host strain XL1 Blue MRF'. Inserts were then subcloned by mass excision using ExAssist helper phage for conversion into phagemid vector pBK-CMV in E. coli host strain XL0R."

ORIGIN

Query Match 45.9%; Score 625; DB 10; Length 653;  
Best Local Similarity 99.5%; Pred. No. 3.8e-104;  
Matches 647; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 67 AAGCTTTAGTGAAGAGATATATGCTGTTGAAGGTCGGAATGCCAAGATGTTA 126  
Db 5 AAGCTTTAGTGAAGAGATATATGCTGTTGAAGGTCGGAATGCCAAGATGTTA 64  
127 CGGAATGATGTTAAACCCATAGTATATCTAAATAACTTGGGATGTTGTTG 186  
65 CGGAATGATGTTAAACCCATAGTATATCTAAATAACTTGGGATGTTGTTG 124  
187 CCGGGGTGCTGCTAACTGAGTGTGAGGCCATGCTCTAGTGTGAAGACAGATTG 246  
Db 125 CCGGGGTGCTGCTAACTGAGTGTGAGGCCATGCTCTAGTGTGAAGACAGATTG 184  
QY 247 GGTATAGTATGATGCTGATGCAAG 306



Db 185 GGTATAGTATGATGCTGATGCAAGAGAGAGGACTTATACACCTGGAAAGAGTGTCC 244  
OY 307 TCATTGAGCCAAAGAGTGTATATCTGCGATTGGATTAGCCTTCATGGCAGCAGCCAGG 366  
Db 245 TCATTGAGCCAAAGAGTGTATATCTGCGATTGGATTAGCCTTCATGGCAGCAGCCAGG 304  
OY 367 GTTACAGCTCATATTAACAATGCTGCTTCTATGAGCTTGAAGAGAAATCATTTCTAT 426  
Db 305 GTTACAGCTCATATTAACAATGCTGCTTCTATGAGCTTGAAGAGAAATCATTTCTAT 364  
OY 427 TAGCTTTGAGCTGAGTGTGTTCTGACAGATCTGCTAAGGGAATGAAAGTGTCTTC 486  
Db 365 TAGCTTTGAGCTGAGTGTGTTCTGACAGATCTGCTAAGGGAATGAAAGTGTCTTC 424  
OY 487 AGAAGGCTGAGAGATGATGCTTAAGACGCCCAATGCTTACATCTTCAATTTGAAA 546  
Db 425 AGAAGGCTGAGAGATGATGCTTAAGACGCCCAATGCTTACATCTTCAATTTGAAA 484  
OY 547 ACCCTGCCAATCCCAAGTTCATTATGAACCACTGCTCCAGAGATATGAAAGGCTCCG 606  
Db 485 ACCCTGCCAATCCCAAGTTCATTATGAACCACTGCTCCAGAGATATGAAAGGCTCCG 544  
OY 607 ATGGAAATTTGATGCAATTTGTTCTGGATAGGCACTGCTGTACATTAACAGTGTCTG 666  
Db 545 ATGGAAATTTGATGCAATTTGTTCTGGATAGGCACTGCTGT - CAATTAACAGTGTCTG 603  
OY 667 GAAATATCTTAAGAGAGAGATCCGATATAAGC-TGATTTGGTGTGA 715  
Db 604 GAAATATCTTAAGAGAGAGATCCGATATAAGC-TGATTTGGTGTGA 653

RESULT 14

CK271210 967 bp mRNA linear EST 12-DEC-2003  
LOCUS EST171288 potato abiotic stress cDNA library Solanum tuberosum cDNA  
DEFINITION clone POACX82 5' end, mRNA sequence.  
CK271210  
VERSION CK271210.1 GI:39828188  
KEYWORDS  
SOURCE Solanum tuberosum (potato)  
ORGANISM  
Solanum tuberosum  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum.

REFERENCE 1 (bases 1 to 967)  
AUTHORS Buell, C.R., Hart, A., Zismann, V., Karameycheva, S.A. and Baker, B.  
TITLE Generation of ESTs from abiotic stressed potato tissue  
JOURNAL Unpublished (2003)  
COMMENT Other ESTs: EST171289  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from TIGR via potato@tigr.org  
Seq primer: ATG TAG ACA CTA TAG.  
Location/Qualifiers

FEATURES  
Source  
1..967  
/organism="Solanum tuberosum"  
/mol type="mRNA"  
/cultivar="Kennebec"  
/db xref="taxon:4113"  
/clone="POACX82"  
/tissue type="abiotic stress treated leaf and root tissue"  
/lab host="DH10B-Tora"  
/clone lib="potato abiotic stress cDNA library"  
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: Solanum tuberosum var. Kennebec plants were  
grown from cuttings on a 16hr light/8 hr dark cycle at 25  
C for 3-4 weeks. Abiotic stress conditions were applied to  
four separate sets of plants. Set 1 involved saturation of  
the soil with 150 mM NaCl and tissues were harvested at  
following application of the salt stress (leaves: 2hr,  
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).

ORIGIN

Query Match 45.0%; Score 612.8; DB 14; Length 967;  
Best Local Similarity 77.4%; Pred. No. 5.3e-102;  
Matches 743; Conservative 0; Mismatches 217; Indels 0; Gaps 0;

Set 2 were grown under the standard conditions and then  
were water stressed by withdrawal of further watering  
applications. Drought stressed plants were harvested after  
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d  
and 5d). Set 3 were grown under the standard conditions  
and then were cold stressed by placement at 4 C. Cold  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d. Set 4 were grown under the standard conditions and  
then were heat stressed by placement at 35 C. Heat  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d and 4d and heat-stressed roots were harvested at 6 hr,  
12 hr, 1 d, and 4d. RNA was isolated from all tissues and  
equal RNA from each tissue and stress was pooled to  
construct the cDNA library. RNA sample."

OY 47. TCATTCTGCTTCTCAAGTCAAGCTTTGTAGTGAAGATTAATGCTGTTGAAGTTC 106  
Db 8 TCATTCTGCTTCTTGTGCAATTCCTAGAAATTAACCTGATTAATGCGGGGAAAGAT 67  
OY 107 CGGAATGCCAAAGATGTTACGGAATGATTTGTAACCAACCCATTAATATCTAAATAA 166  
Db 68 TGGAATGCCAAAGATGTTACGGAATGATTTGTAACCAACCTTGTATACCTGAATTA 127  
OY 167 ACTGCGATGCTGTGTGTTGCCGGGTGCTGCTAACTGAGTGTGATGAGCCATGCTC 226  
Db 128 TGTTGATGATGGGTGTGTGTCACGTGTGCTGCCAAGCTGGAAGCATGAGCCATGCTC 187  
OY 227 TAGTGTGAAGACAGGATTTGGTATATGATTTGCTGATGGAAGAGAGGACTTAT 286  
Db 188 TAGTGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 247  
OY 287 CACACCTGGAAGAAGAGTGTCTCTCATGAGCCCAACAGTGTAACTGCGATTGATTAGC 346  
Db 248 CAAACCCGGCAGAGATGTCCTCATGCAACCTAGAGTGAACCACTGTTGAGATTGGC 307  
OY 347 CTTGATGCGACAGCCAGGCTTACAGCTCATATTAATCAATGCTGCTTCTATGAGTCT 406  
Db 308 ATTCAATGCTGCTGCTTAAGGCTTACAACTCATCATTAAGTCTTCTCAATGAGTCT 367  
OY 407 TGAGAGAGAAATCATTTCTATTAGCTTTTGAGAGCTGAGTGTGTCTGACAGATCTGCTAA 466  
Db 368 TGAGAGAGAAATTAATTTCTGCGTGTCTTCCGCTGAGTGTGTCTTACCGATCCAGCAA 427  
OY 467 GGGAAATGAAGCTGCTGTTCAAGAGCTGAAGATATTGGCTAAAGCCCAATGCTCTA 526  
Db 428 AGGATGAAGGCTGCTTATTCAAGGCTGAAGATTAAGGCCCAACCACTGCTCTA 487  
OY 527 CATACTTCAACAATTTGAAAACCCCTGCCAATCCCAAGTTCATTATGAACCACTGCTTC 586  
Db 488 TATCTTCAACAATTTGAAAACCCCTGCCAATCCCAAGTTCATTATGAACCACTGCTTC 547  
OY 587 AGAGATATGAAGAGCTCCGATGGGAAATTTGATGCAATTTGTTCTGGATAGGCACTGG 646  
Db 548 TGAGATCTGAAGAGCTCCAAATGGGAAATTTGATGCTCTAGTCTCTGGAATTGGAACAG 607  
OY 647 TGGTACATTAACAAGTGTGCTGGAATATCTTAAGAGCAGAAATCCGAATATAAGCTGAT 706  
Db 608 AGGACAGATTAAGTGTGCTGGAATATCTTAAGAGCAGAAATCCGAATATAAGCTGAT 667  
OY 707 TGGTGTGAACCAAGTGAAGTCCAGTCTCTCAGAGAGAAAGCTGTCTCCACACAGAT 766  
Db 668 TGGCGTGAACCAAGTGAAGTCCAGTCTCTCAGAGAGAAAGCTGTCTCCACACAGAT 727  
OY 767 TCAAGGATTTGCTGCTGTTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 826  
Db 728 TCAGGGATTTGCTGCTGTTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 787

QY 827 TGTCAATATCAAGTATGAGCAATAGAACTGCAAGCTTCTTCCCTTAAGAAG 886  
DB 788 AGTTCAGGTTCAAGTGAAGATTCATAGAAATGGCTAAGCTTCTGCAATGAAGAAG 847  
QY 887 CCTATTGTGGGAATATCTTCCGAGCTGAGCTGCTGCTTTCAGATTGCAAAAAG 946  
DB 848 AATGCTAGTGGGAATTCATCTGCTGCTGCGCGCGGCAATTAAGTGTAAAGCG 907  
QY 947 ACCAGAAATGCGGAGCTTATTTGCGCTTTTCCAGCTTGGGAGAGTACT 1006  
DB 908 CCTGAAATGCTGGAGAGCTCATTTGTTATTTTCCAGCTTGGAGAGGATATCT 967

RESULT 15  
BQ630428 613 bp mRNA linear EST 21-OCT-2002  
LOCUS BQ630428.1 GI:21678077  
DEFINITION Gm-c1045-3969 5' similar to TR:081154 081154 CYSTEINE SYNTHASE. ;  
mRNA sequence.

ACCESSION BQ630428.1 GI:21678077  
VERSION BQ630428  
KEYWORDS EST.  
SOURCE Glycine max (soybean)  
ORGANISM Glycine max

REFERENCE 1 (bases 1 to 613)  
AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Bepelding, J., Coryell, V.,  
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,  
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,  
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,  
Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,  
McCann, R., Waterston, R., and Wilson, R.  
Public Soybean EST Project  
Unpublished (1999)

TITLE Public Soybean EST Project  
JOURNAL Contact: Shoemaker R./Public Soybean EST Project  
COMMENT Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130  
South Memorial Parkway Huntville, AL 35801 For further information  
call: (800)-533-4363 or contact: cou@resgen.com web site:  
www.resgen.com

Seq primer: -40RP from Gibco  
High quality sequence stop: 429.

## FEATURES

source

1..613

location/Qualifiers

/organism="Glycine max"

/mol\_type="mRNA"

/db\_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-c1045-3969"

/tissue\_type="Hypocotyl, 9-10 day old etiolated seedlings"

/lab\_host="DH10B"

/clone\_lib="Gm-c1045"

/note="Vector: pBluescriptII SK+; Site 1: EcoRI; Site 2:  
XhoI; This cDNA library was constructed from mRNA isolated  
from etiolated hypocotyl tissue of 9-10 day old seedlings  
of the cultivar Williams 82. Complementary DNA was  
synthesized from mRNA using a primer consisting of a  
poly(dT) primer with a XhoI restriction site. EcoRI  
adapters were ligated to the blunt-ended cDNA fragments  
followed by digestion with EcoRI and XhoI. The cDNA  
fragments were directionally cloned into the EcoRI-XhoI  
restriction site of the pBluescript vector. The ligated  
cDNA fragments were transformed into DH10B host cells  
(Gibco BRL). This library was constructed by Dr. Randy  
Shoemaker."

Query Match 44.5%; Score 606; DB 13; Length 613;  
Best Local Similarity 100.0%; Pred. No. 1.1e-100;  
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 9 GTTGTAGATAGCCGATGCTGCTTCTAGTGTGCTAGCTATTCCTGTTCTCAAGTCAA 68  
DB 8 GTTGTAGATAGCCGATGCTGCTTCTAGTGTGCTAGCTATTCCTGTTCTCAAGTCAA 67  
QY 69 GCTTTGTAGTACGAGATATAATGCTGTTGAAAGGTCGGAATTGCCAAGATGTTACG 128  
DB 68 GCTTTGTAGTACGAGATATAATGCTGTTGAAAGGTCGGAATTGCCAAGATGTTACG 127  
QY 129 GAATGATTTGTAACCCCATTAATATATCTAAATTAATCTGCGATGTTGTTGCC 188  
DB 128 GAATGATTTGTAACCCCATTAATATATCTAAATTAATCTGCGATGTTGTTGCC 187  
QY 189 CGGTTGCTGCTAAACTGAGTGTGATGAGCCATGCTAGTGTGAAGACAGAGATTGGG 248  
DB 188 CGGTTGCTGCTAAACTGAGTGTGATGAGCCATGCTAGTGTGAAGACAGAGATTGGG 247  
QY 249 TATAGTATGATTTGCTGATGAGAGAGAGAGGACTTATCACACTGGAAGAGTCTCTC 308  
DB 248 TATAGTATGATTTGCTGATGAGAGAGAGAGGACTTATCACACTGGAAGAGTCTCTC 307  
QY 309 ATTGAGCCACAAGTGTATATCTGCGATTGATTAGCTTCATGCGACAGCCAGGGGT 368  
DB 308 ATTGAGCCACAAGTGTATATCTGCGATTGATTAGCTTCATGCGACAGCCAGGGGT 367  
QY 369 TACAAGCTCATATTAATGACGCTGCTTCTATGAGTCTGAGAGAGAAATCATTTATTA 428  
DB 368 TACAAGCTCATATTAATGACGCTGCTTCTATGAGTCTGAGAGAGAAATCATTTATTA 427  
QY 429 GCTTTGAGCTGAGTGTGCTTCTGACAGATCTGCTTAAGGAATGAAGGTGCTTCAG 488  
DB 428 GCTTTGAGCTGAGTGTGCTTCTGACAGATCTGCTTAAGGAATGAAGGTGCTTCAG 487  
QY 489 AAGGCTGAAGATATTTGCTAAGACGCCCCAATGCTTACATCTTCAACAATTTGAAAAC 548  
DB 488 AAGGCTGAAGATATTTGCTAAGACGCCCCAATGCTTACATCTTCAACAATTTGAAAAC 547  
QY 549 CTTGCCAATCCCAAGGTTTATATGAACCACTGTGTCAGAGATATGAAAGGCTCCGAT 608  
DB 548 CTTGCCAATCCCAAGGTTTATATGAACCACTGTGTCAGAGATATGAAAGGCTCCGAT 607  
QY 609 GGGAAA 614  
DB 608 GGGAAA 613

Search completed: May 21, 2004, 18:59:53  
Job time : 3965 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2004, 17:34:02 ; Search time 137 Seconds  
(without alignments)  
517.103 Million cell updates/sec

```
Title: US-09-931-457A-30
Perfect score: 1362
Sequence: 1 acctgtgtagtcgtagatag.....aaaaaaaaaaaaaaaaaa 1362
```

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seg:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seg:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seg:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seg:*
5: /cgn2_6/ptodata/2/ina/PTUS.COMB.seg:*
6: /cgn2_6/ptodata/2/ina/backfile1.seg:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	495.8	36.4	975	4	US-09-905-290A-3	Sequence 3, Appl1
2	273	20.0	100848	4	US-09-596-002-39	Sequence 39, Appl1
3	255.4	18.8	1008	4	US-09-543-681A-3561	Sequence 3561, Ap
4	251.2	18.4	1008	4	US-09-328-352-1324	Sequence 1324, Ap
5	251.2	18.4	6115	4	US-08-956-171E-148	Sequence 148, Appl
6	250.2	18.4	10993	4	US-08-961-527-15	Sequence 15, Appl
7	248.2	18.2	927	4	US-09-107-5332A-271	Sequence 271, App
8	245.8	18.0	960	4	US-09-134-001C-493	Sequence 493, App
9	225.2	16.5	4403765	3	US-09-103-840A-2	Sequence 2, Appl1
10	225.2	16.5	4411529	3	US-09-103-840A-1	Sequence 1, Appl1
11	223.8	16.4	1830121	4	US-09-557-884-1	Sequence 1, Appl1
12	223.8	16.4	1830121	4	US-09-643-990A-1	Sequence 1, Appl1
13	218.8	16.1	640681	4	US-09-790-988-1	Sequence 1, Appl1
14	209.8	15.4	927	4	US-09-724-623-10	Sequence 10, Appl1
15	209.8	15.4	8885	4	US-09-634-238-26	Sequence 26, Appl1
16	200.4	14.7	981	4	US-09-489-039A-1995	Sequence 1995, Ap
17	180.6	13.3	2247	4	US-09-252-991A-14920	Sequence 14920, A
18	180.6	13.3	2334	4	US-09-252-991A-14450	Sequence 14450, A
19	167.6	12.3	560	4	US-09-621-976-519	Sequence 519, App
20	147.6	10.8	591	4	US-09-134-000C-121	Sequence 121, App
21	146.6	10.8	912	4	US-09-543-681A-3548	Sequence 3548, Ap
22	142.4	10.5	66986	4	US-09-596-002-29	Sequence 29, Appl
23	139.8	10.3	933	4	US-09-540-236-1422	Sequence 1422, Ap
24	135.4	9.9	924	4	US-09-328-352-4053	Sequence 4053, Ap
25	131.2	9.6	840	4	US-09-134-000C-1996	Sequence 1996, Ap
26	129.4	9.5	927	4	US-09-107-532A-2975	Sequence 2975, Ap
27	124.4	9.1	482	4	US-09-843-472-1	Sequence 1, Appl1

28	120.2	8.8	260	4	US-09-313-294A-3302	Sequence 3302, App
29	116.2	8.5	495	4	US-09-252-991A-15179	Sequence 15179, A
30	110.4	8.1	5014	3	US-09-381-8662-3	Sequence 3, Appli
31	107	7.9	414	4	US-09-134-000C-120	Sequence 120, App
32	95.8	7.0	36800	4	US-08-311-731A-139	Sequence 139, App
33	94.4	6.9	936	4	US-09-252-991A-8456	Sequence 8456, App
34	94	6.9	1524	4	US-10-012-762-14	Sequence 14, Appl
35	94	6.9	1524	4	US-09-704-036B-14	Sequence 14, Appl
36	91.8	6.7	8536	4	US-08-956-171E-278	Sequence 278, App
37	89.6	6.6	2500	3	US-09-318-448-14	Sequence 14, Appl
38	89.6	6.6	2542	1	US-08-120-960-1	Sequence 1, Appli
39	89.6	6.6	2542	4	US-09-347-878-8	Sequence 8, Appli
40	88.2	6.5	939	4	US-09-134-001C-1083	Sequence 1083, App
41	84.6	6.2	921	4	US-09-489-039A-2004	Sequence 2004, App
42	78	5.7	741	4	US-09-252-991A-13320	Sequence 13320, A
43	78	5.7	1149	4	US-09-252-991A-12995	Sequence 12995, A
44	78	5.7	3486	4	US-09-252-991A-12477	Sequence 12477, A
45	77.4	5.7	288	4	US-09-313-294A-5715	Sequence 5715, App

## ALIGNMENTS

```

RESULT 1
US-09-905-290A-3
; Sequence 3, Application US/09905290A
; Patent No. 6605459
; GENERAL INFORMATION:
; APPLICANT: Rice, John
; APPLICANT: Lanning, Beth
; APPLICANT: Crawford, John
; APPLICANT: Nye, Gordon
; TITLE OF INVENTION: METHODS FOR MEASURING CYSTEINE AND DETERMINING CYSTEINE SYNTHASE
; FILE REFERENCE: Docket No. 66054592090US
; CURRENT APPLICATION NUMBER: US/09/905,290A
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 975
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURES:
; NAME/KEY: misc_feature
; OTHER INFORMATION: AJ011603 GenBank
US-09-905-290A-3

```

Query Match	36.4%;	Score 495.8;	DB 4;	Length 975;
Best Local Similarity	69.4%;	Pred. No. 3.4e-132;		
Matches 674; Conservative	0;	Mismatches 297;	Indels 0;	Gaps 0;

[illegible]



```

Db 301 CCTCATCATGAGCCTTGAGAGAAATCATTTTACATTAGTGGGAGGTTAC 360
QY 450 CTGACAGATCTGCTTAAGAGATGAAAGTCTGCTTCAAGAGCTGAAGATATTGGCT 509
Db 361 CTACAGATCTGCTTAAGAGATGAAAGTCTGCTTCAAGAGCTGAAGATATTGGCT 420
QY 510 AAGAGCCCAATGCTTACATCTTCAACATTTGAAAACCTGCGCAATCCCAAGCTTCA 569
Db 421 AAAATCCAGATAGTATCATCTAGAACAGCTTCAAAAATCTTCAAAACCCCAACTCAT 480
QY 570 TATGAACCACTGCTTCAAGATATGAAAAGCTCCGATGGGAAATTTGATGATTGTT 629
Db 481 TATGAACCACTGCTTCAAGATATGAAAAGCTCCGATGGGAAATTTGATGATTGTT 540
QY 630 TCTGGATAGGCACTGCTTCAAGATATGAAAAGCTCCGATGGGAAATTTGATGATTGTT 689
Db 541 GCGGCTGTTGAAGCTGCTTCAAGATATGAAAAGCTCCGATGGGAAATTTGATGATTGTT 600
QY 690 CCGAATATTAAGCTGCTTCAAGATATGAAAAGCTCCGATGGGAAATTTGATGATTGTT 749
Db 601 AAAGACTTAAGCTGCTTCAAGATATGAAAAGCTCCGATGGGAAATTTGATGATTGTT 660
QY 750 CCGGCTGTTGAAGCTGCTTCAAGATATGAAAAGCTCCGATGGGAAATTTGATGATTGTT 809
Db 661 CCGGCTGTTGAAGCTGCTTCAAGATATGAAAAGCTCCGATGGGAAATTTGATGATTGTT 720
QY 810 AATCTTCTGATGAGTGTTCATATATCAAGTGAAGCAATGAAATGCAAAAGCTT 869
Db 721 AAGCTTCTGATGAGTGTTCATATATCAAGTGAAGCAATGAAATGCAAAAGCTT 780
QY 870 CTGCGCTTAAGAGCTTATTTGAGGAAATATCTCCGAGCTGAGCTGCTGCTGCT 929
Db 781 CTGCGCTTAAGAGCTTATTTGAGGAAATATCTCCGAGCTGAGCTGCTGCTGCTGCT 840
QY 930 TTTCAGATGCAAAAGCAAGCAAAATGCGGAGCTTATTTGCGGTTTTCAGAGC 989
Db 841 ATTAAGTGGCAAGCGGCAAGCAAAATGCGGAGCTTATTTGCGGTTTTCAGAGC 900
QY 990 TTGCGGAGAGCTGCTTCTCCGCTATTTGAGTCAAGTGAAGCAAGCTGAAAGC 1049
Db 901 GGTGGGAGAGCTGCTTCTCCGCTATTTGAGTCAAGTGAAGCAAGCTGAAAGC 960
QY 1050 ATGACTTTTGA 1060
Db 961 TTGCCAATTCA 971

```

```

RESULT 2
US-09-596-002-39
; Sequence 39, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; PRIOR FILING DATE: 2000-06-16
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 39
; LENGTH: 100848
; TYPE: DNA
; ORGANISM: M. catarrhalis
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte template ID No. 6632636 39
; PUBLICATION INFORMATION:
; US-09-596-002-39

```

```

Query Match 20.0%; Score 273; DB 4; Length 100848;
Best Local Similarity 57.7%; Pred. No. 3.1e-67;
Matches 527; Conservative 0; Mismatches 380; Indels 6; Gaps 2;
QY 111 ATTGCCAAGATGTTACGGAATTTGTTAAACCCCATTAATATCTAATAAAGCTT 170
Db 17953 ATTGCCAAGATGTTACGGAATTTGTTAAACCCCATTAATATCTAATAAAGCTT 18012
QY 171 GCGGATGTTGTTGCTCCCGGCTGCTGCTAACTGAGTTGATGAGCCATGCTTACT 230
Db 18013 ACCGAAGCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 18072
QY 231 GTGAAGCAGGATTTGGTATGATGATGATGATGATGATGATGATGATGATGATGATG 290
Db 18073 GTGAAGCAGGATTTGGTATGATGATGATGATGATGATGATGATGATGATGATGATG 18130
QY 291 CCGGAAAGAGTGTCTCTATGAGCAACCAAGTGTATATCTGCAATTTGATGATGATG 350
Db 18131 -CCGAAGCAGGATTTGGTATGATGATGATGATGATGATGATGATGATGATGATGATG 18189
QY 351 ATGCGAGCAGCAGGAGGTTACAGCTCATATTAATTAATTAATTAATTAATTAATTA 410
Db 18190 ATGCGAGCAGCAGGAGGTTACAGCTCATATTAATTAATTAATTAATTAATTAATTA 18249
QY 411 AGAAGATCATTTATTTAGCTTTTGAAGCTGATGATGATGATGATGATGATGATGATG 470
Db 18250 AGAAGATCATTTATTTAGCTTTTGAAGCTGATGATGATGATGATGATGATGATGATG 18309
QY 471 ATGAAGTGTCTGTTTCAAGAGCTGAAGATATTTGCTTAAGACGCTTATGCTTACATA 530
Db 18310 ATGAAGTGTCTGTTTCAAGAGCTGAAGATATTTGCTTAAGACGCTTATGCTTACATA 18366
QY 531 CTTCACAAATTTGAAAACCTGCAATCCCAAGCTTCAATTAATTAATTAATTAATTA 590
Db 18367 CTTCACAAATTTGAAAACCTGCAATCCCAAGCTTCAATTAATTAATTAATTAATTA 18426
QY 591 ATATGAAGAGCTCCGATGGAATTTGATGATGATGATGATGATGATGATGATGATGAT 650
Db 18427 ATATGAAGAGCTCCGATGGAATTTGATGATGATGATGATGATGATGATGATGATGAT 18486
QY 651 ACAATTAAGAGCTGCTGAAAATTTTAAAGAGCAGATCCGATATTAAGCTGATTTGT 710
Db 18487 ACAATTAAGAGCTGCTGAAAATTTTAAAGAGCAGATCCGATATTAAGCTGATTTGT 18546
QY 711 GTGAACAGTTGAAAGTCCAGTGTCTCAGAGAGAAAGCTGCTCACAACAAGATTCAA 770
Db 18547 GTGAACAGTTGAAAGTCCAGTGTCTCAGAGAGAAAGCTGCTCACAACAAGATTCAA 18606
QY 771 GGGATGCTGCTGTTTATCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 830
Db 18607 GGGATGCTGCTGTTTATCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 18666
QY 831 CAATATCAAGTATGAAGCAATGAAGCAATGAAGCAATGAAGCAATGAAGCAATGAAG 890
Db 18667 CAATATCAAGTATGAAGCAATGAAGCAATGAAGCAATGAAGCAATGAAGCAATGAAG 18726
QY 891 TTGTTGGAATATCTCCGAGCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 950
Db 18727 TTGTTGGAATATCTCCGAGCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 18786
QY 951 GAAATGCGGGAAGCTTATTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1010
Db 18787 GAAATGCGGGAAGCTTATTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 18846
QY 1011 TCCGCTGCTTGT 1023
Db 18847 TCCGCTGCTTGT 18859

```

```

RESULT 3
US-09-543-681A-3561
; Sequence 3561, Application US/09543681A
; Patent No. 6605709

```

```

; GENERAL INFORMATION:
; APPLICANT: GARY BRETTON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 3561
; LENGTH: 1008
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-3561

```

Query Match	18.8%;	Score 255.4;	DB 4;	Length 1008;
Best Local Similarity	58.3%;	Pred. No. 2.8e-63;		
Matches 561;	Conservative	0;	Mismatches 366;	Indels 36;
			Gaps	5,

QY	85	ATATATATGCGCTGTTGAAGGTCCGGAATTGCCAAGATGTTACGGAATTGATTGGTAAA	144
Db	38	ATGTACCGAGGATAAAAATGACCAAAATTTTCGAAGATTAATCCGCAAACTATTGGCCATA	97
QY	145	CCCCATTAGTATATCTTAATAAATACTTCCGAGATGTTGTCTTCCCGGTTGCTGCTAAAC	204
Db	98	CCCCCCTAGTTGCTTTAAACATTTCG-----GTAAATGCAATATCTTAGCTAAAG	148
QY	205	TGGAGTTGATGAGCCATGCTCTAGTGTGAAGACACAGATTGGGTATAGTATGATTGCTG	264
Db	149	TAGAAATCTCGTAATCCAAAGCTTTAGTGTAAATATGCCGTATCCGTGCAAAATATGATTGGG	208
QY	265	ATGCAGAAGAGAAGGAGCTTATACACACCTGGAAGAGTGTCTCATTGAGCCAACTAGTG	324
Db	209	ATGCAGAAAAAAGGTATCTGAAACCCGGCGTA--GAGCTAGTGAAGCCCAACCACTG	265
QY	325	GTAACTACTGGCATTGGATTAGCCTTCATGCGACGACGACGCGGTACAGCTCATATTA	384
Db	266	GTAACTACTGGGATTGCTTGCTTATGTAGCTGCTGCGGTGGCTATTAAGCTGACACTGA	325
QY	385	CAATGCTGCTTCTATGAGTCTTGAGAGAAGAATCATTTCTATTAGCTTTTGAGCTGAGT	444
Db	326	CTATGCCTGATACCATGAGTATGAGCGCGTAATTATTAAGCAATTAGTGTAAAC	385
QY	445	TGCTTCTGACAGATCCTGCTAAGGAATGAAGGTGCTGTTCAGAAAGCTGAAGAGATAT	504
Db	386	TTGTGTTAAGTGAAGGCGGTAAAGTATGAAGGAGCGATGATAAAGCGAAGAAATTC	445
QY	505	TGCTAAGACGCCCAATGCTTAC--ATACTTCAACAATTTGAAAACCTGCCAATCCCA	561
Db	446	GTGATAGCGATCCTAGCCGTTACCTTCTACTACACAATTTAGTAATCCAGCAAAACCCAG	505
QY	562	AGGTTCAATTATGAACCCACTGTCACAGATATGGAAGGCTCCGATGGGAAATTTGATG	621
Db	506	AAATTCATGAGAAAAACCAAGGCTTGAATTTGGAATGACACTGATGGCAAGTTGATG	565
QY	622	CATTGTTTCTGGATAGGCACTGCTGTACAAATAACAGGTGCTGGAATAATATCTTAAG	681
Db	566	TAGTTATCGCAGGTGTGGAACGGGTGTACCATCAAGGTATTTGCTGCTATTGAAAA	625
QY	682	A---GCAGAAATCCGAATATAAAGCTGATTGTTGTGGAACCAAGTTGAAGTCCAGTGTCT	738
Db	626	ATACTCAAGGTAAAGCTGTACCAATGTTGCGGTAGAACCTTAAGATTTCTCTGTATTTA	685
QY	739	CAGGAGGA-----AAGCTGTGTCACACAAGATTCAAGGGAATTGTTG	780
Db	686	GCCAAAGCATTTAGCAGGTGAAGAAATTAAGCTGTGTCACATTAATAATCCAAAGTATTGTTG	745
QY	781	CTGCTTTATCCCTGCTGTCTTGAGATCAATCTTCTTGATGAAGTTGTTCAATATCAA	840
Db	746	CAGTTTATCCCTGAAGAACTTAGAGTTATCACTATTAGATGCGGTGATCCAAATTACTA	805
QY	841	GTGATGAACGATAGAACTGCAAGCTTCTTGCGCTTAAAGAGGCTTATTGTGGAA	900

Db 806 ACGAAGAAGCATTGGAACAGCTCGTGAAGTATGACCAAGAAGTATTTAGCGGTA 865

QY 901 TATCTTCCGAGCTGCAGCTGCTGCTTTTCAATTCAAAAAGACCAGAAATGCCG 960

Db 866 TTTCATCCGTCGCGCAATGCGGCTGCAGTGAATTAAGCTAAAGAGCCAGAATTTGCTA 925

QY 961 GGAAGCTATTTGTCGCGTTTTCACAGCTTCGGGGAAGTAACTGTCTCCGTCTAT 1020

Db 926 ATAAGGAATCGTAGTATTTAACCTTCTCAGGTGAACGTTATCTCAGCACCCACCTTT 985

QY 1021 TTG 1023

Db 986 TTG 988

RESULT 4  
US-09-328-352-1324  
; Sequence 1324, Application US/09328352

```

: GENERAL INFORMATION:
: APPLICANT: Gary L. Breton et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
: TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: GTC99-03PA
: CURRENT APPLICATION NUMBER: US/09/328,352
: CURRENT FILING DATE: 1999-06-04
: NUMBER OF SEQ ID NOS: 8252
: SEQ ID NO 1324
: LENGTH: 1008
: TYPE: DNA
: ORGANISM: Acinetobacter baumannii
US-09-328-352-1324

```

Query Match	18.4%;	Score 251.2;	DB 4;	length 1008;
Best Local Similarity	58.3%;	Pred. No. 4.4e-62;		
Matches 549;	Conservative	0;	Mismatches 363;	Indels 30;
				Gaps 5;

D	b		CGAGTTATTCAATAATGACAGAAGCTATTGGTAACTCCTTAGTA---CGTATTAA	110
O	y	107	CGGAATTCGCAAAAGATGTACGGAAATGATTTGGTAAAAACCCCATAGATATCTAATTA	108
D	b	54	CGAGTTATTCAATAATGACAGAAGCTATTGGTAACTCCTTAGTA---CGTATTAA	110
O	y	167	ACTTGCGATGGTTGTGTGCCCGGCTGCTGCTAAACTGGAAGTTGATGGAGCCATGCTC	226
D	b	111	TGCTTGATTAAACTGATGCACTGTATTAGCCAAAGTAGAAAGTCGTACCAGCTTT	170
O	y	227	TACTGTGAAGCACAGGATTGGGTATAGTATGATGTGCTGATGCAGAAAGAGGACTTAT	286
D	b	171	TTCTGTGAATGCCGTATGGAGCTGCGCTGATTGCAGATGCAGAAAAAGAGCGCGTT	230
O	y	287	CACACTGGAAGAAGSTGTCTCTCATTTGAGCCACAAGAAGTGTATTACTGGCATTGATTAGC	346
D	b	231	GAAAGAAAGCATGCATATTG---TTGAACCAACCAAGTGTATTACTGTATTGCTTGGC	287
O	y	347	CTTCATGGCAGCAGCCAGGGGTACAAGCTCATATTACAAATGCTGCTTCTATGAGTCT	406
D	b	288	GTTCTGAGCGCGGCAAAAAGTTATTCCATTACTTAACTATGCTGCACAAGCATGAGCCT	347
O	y	407	TGAGAGAAGAATCATTTCTATTAGCTTTTGGAGCTGAGTTGGTTCTGCACAGATCCTGCTAA	466
D	b	348	TGAACGTAGAAAAGTGTTAAGAAGCACTCGGTGCCAATTTAGTCTTAACAGAACCAAGCAA	407
O	y	467	GGGAATGAAAAGTGCTGTTTCAGAAGGCTGAAGAGATATTGGCTAAGACGCCCAATGCTTA	526
D	b	408	GGGAATGAAAAGTGCAGTTGATGAAGCTGTACGATTAGCAACTGAAACAGCCAGAAGTTTA	467
O	y	527	---CATACTTCAACAATTTGAAAAACCTGCCAATCCCAGGTTCAATTAGAAAACCACTGG	583
D	b	468	TTTTCTGCCACAACAGTTTGAAAACCCAGCTAATCCACAATATCAATGTGATACGACTGG	527
O	y	584	TCCAGAGATATGAAAAGGCTCCGATGGGAAAAATTGATGCATTTGTTCTGGGATAGGCAC	643
D	b	528	ACCAGAAATCTGGCAAGCAACAGGTGGGCAAGTTGATATCTTGTAAGCTGCTGTAGGTAC	587

```

QY 644 TGGTGTACATTAACAGGTGCTGMAAATATCTTA---AAGAGCAGAAATCCGAATATPAA 700
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 588 AGAGAGTACGATTAACGGGTATTTCTCGTATTTTGAACAAGTTGAGAAATAGCCATTGTA 647
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 701 GCTGATGTGTGGAACCAAGTTGAAGTCCAGTCTCTCAGAGAGAAAGCTGT----- 755
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 648 CTCAGTAGCAGTTGAACTGTGAGTCTCAATTAATTAATCAAAACAAAATGCTGAGAA 707
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 756 -----CCACACAGATTCAAGGAGTGTGCTGTTTATCCCTGCTGCTT 802
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 708 CATTACACCCGACCGCAATAAATTCAGGAATGAGCAAAATTTATCCAAAGAAATCT 767
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 803 GGAAGTCAATCTTCTGATGAAGTTGTTCAAAATCAAGTATGAGCAATAGAAACTGC 862
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 768 TGATTTAGATTGTTGATGAAGTCTTACAGTAGTAGTGAAGAGCAATTCATGAGGC 827
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 863 AAAGCTTCTGCGCTTAAGAGAGCTTATTTGAGGAATATCTCCGAGCTGACAGCTGC 922
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 828 GAGAAAGTGTGCCACCCAGAGAGTATTTAGTGGTATTTCAAGTGTGCTGCATTGGC 887
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 923 TGCTGCTTTTCAAGTTGCAAAAAGACCAAGAAATGCCGGAAGCTTATTTGCTGCTTT 982
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 888 AGCAGTGGCTAAATTCAGAAAGCTCTGAAATGAGAGTAAACAAATGTTGTAATCT 947
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 983 TCCCAAGCTTGGGAGAGTACCTGCTCCGCTGCTATTGA 1024
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 948 TCCAGATAGTGGGAAACCTATTATCTCTGTTTATTGA 989
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 5  
US-08-956-171E-148  
; Sequence 148, Application US/08956171E  
; Patent No. 6593114

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

Gil H. Choi

Patrick S. Dillon

Craig A. Rosen

Steven C. Barash

Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5256

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/956,171E

FILING DATE: 20-Oct-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/009,861

FILING DATE: January 5, 1996

APPLICATION NUMBER: 08/781,986

FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Mark J. Hyman

REGISTRATION NUMBER: 46,789

REFERENCE/DOCKET NUMBER: PB248P1

TELEPHONE: (240) 314-1224

TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 148:

SEQUENCE CHARACTERISTICS:

```

; LENGTH: 6115 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 148:
US-08-956-171E-148

```

```

Query Match 18.4%; Score 251.2; DB 4; Length 6115;
Best Local Similarity 58.7%; Pred. No. 1.2e-61;
Matches 493; Conservative 0; Mismatches 338; Indels 9; Gaps 3;

```

```

QY 121 ATGTACCGAATTGATGTGTAACCCCATTAATATCTAATAAATGCGATGTT 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4760 ATATTACTCAATTAATTGCGCGGTACACCGTAGTCAATGAGAAATGATAGATGACA 4819
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 GTGTTGCGCGGCTGCTGCTAACTGAGTGTGATGAGCCATGCTCTAGTGAAGACA 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4820 ATGACGACAGTGTATTATGTAATAATTGGAATATCAAAATCCAGTGTCTCTGAAGATA 4879
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 GGAATGCGTATGATGATGCTGCTGATGCAAGAGAGAGAGGACTTATCACACCTGAAAGA 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4880 GAATGCTTTAGCAATGATTTGAAATAAGCAGAGCGAGAGAGGAAATTAACCTGCGCA-- 4937
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 301 GTGTCTCATGAGCCCAACAGTGTGTAATCTGCAATTGGAATGCTTATGCGCAGAG 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4938 -TACAAATGTAAGACCAACAGTGTGTAATCAAGTATCGGTTTATGATTTGATGCTG 4996
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 361 CCAGGGGTACAAAGCTCATTAATTAATGCTGCTCTTATGAGTCTTGAGAGAGAAATCA 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4997 CTAAAGATATTAAGCAGTATTTACTTATGCCCCGAAACAAATGAGCCAAAGAGCTGTAAT 5056
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 421 TTCTATTACCTTTTGGAGCTGAGTGTGTTCTGACAGATCTCTGTAAGGAATGAAAGTG 480
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5057 TATTAAAGCATACGAGTGCAGAAATTAATTTTAAAGCTGATCAGAAAGGATG 5116
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 481 CTGTTCAAGAGCTGAAGATATTTGGCTAAGAGCGCCCAATGCTTACATTAATCAACAT 540
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5117 CAATTAAAGAACTTAAGA--ATTGAAGAGAGAAACATGTTACTTCCAGCCCAACAAT 5173
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 541 TTGAAACCTGCGCAATCCCAAGGTTCAATATGAACCACTGCTCAGAGATATGGAAG 600
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5174 TTGAAACCTGCGCAACCTCGAAGTCAATGATTAATCAAGGTTATTACAAAC 5233
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 601 GCTCGATGGGAAATTT--GATGATTTGTTCTGGATAGGCACTGTTGTAATTA 657
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5234 AATTGAAGGAAACTATGATGCTGCTTCTAGCTGTTGTTGTTGTTGTTGTTGTTAT 5293
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 658 CAGTGTGGAATAATATCTTAAGAGCAGAAATCCGAATATTAAGCTGATGTTGTTGAAC 717
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5294 CTGCTGATGTAAGTCTGTAAGAAAGATATCTTAACATGCAATGTTGTTGTTATAGAC 5353
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 718 CAGTTGAAGTCCAGTGTCTCTCAGAGAGAAAGCTGCTGTCACACAGATTCAAGGATG 777
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5354 CTGAGGCTTCTCAGATTAAGAGGCTGAGGAGCCAGGTCACATTAATTAAGAGTTAG 5413
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 778 GTGCTGTTTATCCCTGCTGCTGTTGGAAGTCAATCTTCTGATGAAGTGTCAATAT 837
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5414 GTGCTGATTTATTCAGGCACTTGAATACAGAAATCTATGACAGTATTAATAAGTAG 5473
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 838 CAAGTATGAAGCAATAGAAACTGCAAAAGCTTCTTCCGCTTAAGAGGCTATTTGTG 897
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5474 GAATGATACAGCATGAATATGCTGCTGAGTTGCTAAGAGGAGTATTTAGCAG 5533
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 898 GAATATCTCCGAGCTGACAGTGTGCTGCTTTTCAAGATTGCAAAAAGCAGAAATG 957
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5534 GTATTTCATCAGGTGCTGCAATTTATGCTGCAATCAAAAAGCAAAAAGATTAGGAAAG 5593
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 6

US-08-961-527-15

; Sequence 15, Application US/08961527

; Patent No. 6420135

; GENERAL INFORMATION:



APPLICANT: Charles Kunsch  
 TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
 NUMBER OF SEQUENCES: 391  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Human Genome Sciences, Inc.  
 STREET: 9410 Key West Avenue  
 CITY: Rockville  
 STATE: Maryland  
 COUNTRY: USA  
 ZIP: 20850  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
 COMPUTER: HP Vectra 486/33  
 OPERATING SYSTEM: MSDOS version 6.2  
 SOFTWARE: ASCII Text  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/961,527  
 FILING DATE:  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Brookes, A. Anders  
 REGISTRATION NUMBER: 36,373  
 REFERENCE/DOCKET NUMBER: PB340P1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (301) 309-8504  
 TELEFAX: (301) 309-8512  
 INFORMATION FOR SEQ ID NO: 15:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 10993 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 US-08-961-527-15

Query Match 18.4%; Score 250.2; DB 4; length 10993;  
 Best Local Similarity 57.6%; Pred. No. 3.2e-61;  
 Matches 529; Conservative 0; Mismatches 378; Indels 12; Gaps 4;

QY 111 ATTGCCAAGATGTTACGGAATGATGTAACCCCAATTAGTATATCTAATAACTT 170  
 DB 7400 ATTATACAACATTACTGAAATTAATGCTCAACACCGAATGTTAACTTAACAACATC 7459  
 QY 171 GCGGATGTTGTTGTCGCCGGTTCCTGCTAACTGAGTTGATGAGCCATGCTTACT 230  
 DB 7460 GTGCCAGAGGTGCTGCAGACGCTTAATAAGCTTGAAGCAATTATCTGTTCACT 7519  
 QY 231 GTGAAGACAGATTGGGTATAGTATGCTGATGAGAGAGAGAGAGAGAGAGAGAGAG 290  
 DB 7520 GTAAAGACGATATGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7579  
 QY 291 CCTGGAAGAGTGTCTCTATGAGCAACAAGTGTATATGCAATGCAATTGAGCTTTC 350  
 DB 7580 CCTGTTTACT--ATTGTTGAAGCAACAAGTGTATATGCAATGCAATTGAGCTTTC 7636  
 QY 351 ATGGCAGACGACGAGGTTTACAGCTTATATTAATGAGCTGCTTCTATGAGCTTTCAG 410  
 DB 7637 GTAGTGTCTGTAAGAGGATTAAGCTGATGCTGATGCTGTAAGAGAGAGAGAGAG 7696  
 QY 411 AGAAGATTCATTCTATTAGCTTTGAGCTGAGTGTGTTCTGACAGATCTGCTAAGGA 470  
 DB 7697 CGACGTAAATATATCCAGCTTATGCTGTAAGCTGCTGTAAGAGAGAGAGAGAGAG 7756  
 QY 471 ATGAAGGTGCTGTTTCAAGAGGCTGAAGAGATATGCTAAGAGAGAGAGAGAGAGAG 530  
 DB 7757 ATGAAGGTGCTTATTGCTAAGGCTCAAGAAATCGCTGCTGA--AGCTGATGCTTCTT 7813  
 QY 531 CTCAACATTTGAAACCTGCGCAATCCCAAGCTTATATGAACCACTGCTCAGAG 590  
 DB 7814 CCTCTCAATTTGACATCCAGCTTATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7873

QY 591 ATATGGAAGGCTCCG--ATGGGAAATATGATTCATTGTTTCTGGGATAGGCACTGCT 647  
 DB 7874 ATACTAGCTGCTTCCGTAAGATGATGATGATGCTTGTGCTGAGTAGTACTGCT 7933  
 QY 648 GGTACATTAACAGGTCTGGAATAATATCTTAAGAGAGAGAGAGAGAGAGAGAGAGAG 707  
 DB 7934 GGAACGATTTCTGCTGTTCTCATGCACTCAATCAGAAATTTCTAATCAATTTT 7993  
 QY 708 GGTGTGAACAGTGAAGAGTCCAGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 767  
 DB 7994 GCACTAGAGAGAGATGAATCTGCTATTTATCTGCTGAGAGAGAGAGAGAGAGAGAT 8053  
 QY 768 CAAGGATGCTGCTGTTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 827  
 DB 8054 CAAGGATCTGAGTGAATTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8113  
 QY 828 GTCAATATCAAGTGAAGCAATAGAACTGCAAAAGCTTCTGCTGCTTAAAGAGAGAG 887  
 DB 8114 GTTCGTGAACATGATGAGCTCTGCACTGCGAGCTGAATTTGTTGAAGAGAGAGAG 8173  
 QY 888 CATTGTTGGAATATCTTCCGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 947  
 DB 8174 TTCTTGTAGGATTTCTCAGCTGAGCTATCTAAGAGAGAGAGAGAGAGAGAGAGAGAG 8233  
 QY 948 CCAAGAAATGCCGGAAGCTTATGTTGCTGCTTTCACAGCTTGGGAGAGAGAGAGAGAG 1007  
 DB 8234 TTAG--GTACAGGTAAAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8290  
 QY 1008 TCCTCCGCTGCTATTTGACT 1026  
 DB 8291 TCTACAGCACTTATGAAT 8309

RESULT 7  
 US-09-107-532A-271  
 Sequence 271, Application US/09107532A  
 Patent No. 6583275  
 GENERAL INFORMATION:  
 APPLICANT: Lynn A Doucette-Stamm and David Bush  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
 NUMBER OF SEQUENCES: 7310  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
 STREET: 100 Beaver Street  
 CITY: Waltham  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02354  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: CD-ROM ISO9660  
 COMPUTER: PC  
 OPERATING SYSTEM: <Unknown>  
 SOFTWARE: ASCII  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/107,532A  
 FILING DATE: 30-Jun-1998  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/085,598  
 FILING DATE: 14 May 1998  
 APPLICATION NUMBER: 60/051571  
 FILING DATE: July 2, 1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Ariniello, Pamela Deneke  
 REGISTRATION NUMBER: 40,489  
 REFERENCE/DOCKET NUMBER: GTC-012  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (781)893-5007  
 TELEFAX: (781)893-8277  
 INFORMATION FOR SEQ ID NO: 271:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 927 base pairs  
 TYPE: nucleic acid

STRANDEDNESS: double  
 TOPOLOGY: circular  
 MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Enterococcus faecium  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (B) LOCATION 1..927  
 SEQUENCE DESCRIPTION: SEQ ID NO: 271:  
 US-09-107-532A-271

Query Match 18.2%; Score 248.2; DB 4; Length 927;  
 Best Local Similarity 58.0%; Pred. No. 3.1e-61;  
 Matches 499; Conservative 0; Mismatches 353; Indels 9; Gaps 3;

QY 123 GTTACGGAATTGATGGTAAACCCCATTTAGTATCTAATTAACCTGGGATGTTGT 182  
 DB 22 GTACCGAGTGTGATGGCAAAACCCCATTCCTCAATTAATTAATCGTACCAAGAT 81  
 QY 183 GTTCCCGGGTGTCTGCTAACTGAGTTGATGAGCCATGCTCTAGTGTGAAGACAGG 242  
 DB 82 TCAGCAGATGTTTGTAAAGCTTGAAGTTTAAATCCAGAGGAAGTGTAAAGATCGT 141  
 QY 243 ATTGGGTATAGTATGATGCTGATGACAGAGAGAGGACTTATACACCTGGAGAAAGT 302  
 DB 142 ATTGCTTAAATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 201  
 QY 303 GTCTCTATGAGCCCAAGTGTATATGCTGATGAGTTAGCTTCTGATGAGAGAGAG 362  
 DB 202 ---ATCATTGAACCTTCTCTGGAATACAGAGATCGGATTTGCGATGAGAGAGTTC 258  
 QY 363 AGGGGTACAGCTCATTAATTAATGATGCTGCTCTGATGAGTCTTGAGAGAGATCAT 422  
 DB 259 AAGGATATTAAGTATCATGCTGATGCTGAGAGAGATGATGATGAGAGAGAGAGAG 318  
 QY 423 CTATTAGCTTTGAGAGTGTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 482  
 DB 319 ATGAAAGGATGAG 378  
 QY 483 GTTCAGAGAGCTGAGAGATGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 542  
 DB 379 AT---CAG 435  
 QY 543 GAAAGCCCTGCCAATCCCAAGTCTTATTAAGAGAGAGAGAGAGAGAGAGAGAGAG 602  
 DB 436 GAAAG 495  
 QY 603 TCCGATGGAA---AATGATGATGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 659  
 DB 496 TTTGAGTAAATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 555  
 QY 660 GGTGCTGGAATATATCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 719  
 DB 556 GGGCTGGCCGAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 615  
 QY 720 GTTGAAGAGTCCAGTCTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 779  
 DB 616 GCAAGATCTGCTATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 675  
 QY 780 GCTGCTTTTATCCCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 839  
 DB 676 ACAAGTCTTGTCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 735  
 QY 840 AGTGTGAAG 899  
 DB 736 GAG 795  
 QY 900 ATATCTTCCGAG 959  
 DB 796 ATCTCATCTGCGGCTGCAATGCTGCGGCTTAAAGAGAGAGAGAGAGAGAGAGAG 855

QY 960 GGGAGCTTATGTTGCCGTT 980  
 DB 856 AAGAGGATTAAGCTGTGCTT 876

RESULT 8  
 US-09-134-001C-493  
 ; Sequence 493, Application US/09134001C  
 ; Patent No. 6380370  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
 ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: GTC-007  
 ; CURRENT APPLICATION NUMBER: US/09/134,001C  
 ; PRIOR FILING DATE: 1998-08-13  
 ; PRIOR APPLICATION NUMBER: US 60/064,964  
 ; PRIOR FILING DATE: 1997-11-08  
 ; PRIOR APPLICATION NUMBER: US 60/055,779  
 ; PRIOR FILING DATE: 1997-08-14  
 ; NUMBER OF SEQ ID NOS: 5674  
 ; SEQ ID NO 493  
 ; LENGTH: 960  
 ; TYPE: DNA  
 ; ORGANISM: Staphylococcus epidermidis  
 ; US-09-134-001C-493

Query Match 18.0%; Score 245.8; DB 4; Length 960;  
 Best Local Similarity 57.6%; Pred. No. 1.5e-60;  
 Matches 521; Conservative 0; Mismatches 372; Indels 12; Gaps 4;

QY 121 ATGTTACGGAATTGATGGTAAACCCCATTTAGTATCTAATTAACCTGGGATGTT 180  
 DB 50 ATGTTACGGAATTGATGGTAAACCCCATTTAGTATCTAATTAACCTGGGATGTT 109  
 QY 181 GTTCCCGGGTGTCTGCTAACTGAGTTGATGAGCCATGCTCTAGTGTGAAGAGCA 240  
 DB 110 ATGACGCTGATATTATGTTAGTGAATATCAAAATCCAGTGTGTTGGTAAAGATC 169  
 QY 241 GGATGAGTATGATGATGATGCTGATGACAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
 DB 170 GTATGCTTTAGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 227  
 QY 301 GTGCTCTCATGAGCCCAAGTGTATTAATGATGATGATGATGATGATGATGATGATG 360  
 DB 228 -TACATGCTGAGCCCTACGAGTGTGATGATGATGATGATGATGATGATGATGATG 286  
 QY 361 CCAGGGGTACAGCTCATTAATTAATGATGATGATGATGATGATGATGATGATGATG 420  
 DB 287 CCAAGGGGTACAGCTCATTAATTAATGATGATGATGATGATGATGATGATGATGATG 346  
 QY 421 TTTTATGAGCTTTTGGAGCTGAGTGTGCTGACAGATCTGCTTAAGGAGATGAAAGTG 480  
 DB 347 TATTAAGCTTATGCTGCTGAGTGTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 406  
 QY 481 CTGTTCAAGAGCTGAGAGATGATGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540  
 DB 407 CAATTAAG 463  
 QY 541 TTGAAAGCCCTGCCAATCCCAAGTGTATTAATGAGAGAGAGAGAGAGAGAGAGAGAG 600  
 DB 464 TCGAAAGCCCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 523  
 QY 601 GCTCCGA---TGGGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 657  
 DB 524 AATTTGAAGGTCGACAAATGATGATGATGATGATGATGATGATGATGATGATGATG 583  
 QY 658 CAGTGTGAG 717  
 DB 584 CTGTTGTTGTAAGTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 643  
 QY 718 CAGTGAAGTCCAGTCTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 777

Mon May 24 08:18:49 2004

us-09-931-457a-30.cni

**Page 7**

Db	644	CTGAAGCTTCTCCAGTATTAAAGCGGTGGTGAACCAAGGCCCTCATAAATTACAAGGATTGG	703
Qy	778	GTGCTGGTTTATCCCTGGTGTCTTGGCAAGTCAATCTTCTTGATGAAGTTGTTCAAAATAT	837
Db	704	GAGCAGGTTTGGTACTGTAATCTTTAAATACAGAAAGTTTATGACAGCATCATCAAAAGTAG	763
Qy	838	CAAGTGATGAAGCAATAGAACTGCCAAAGCTTCTTGCGCTTTAAAGAAGGCTATTGTGG	897
Db	764	GTAATGATTAAGTGTATGATGATAGGACAGCTGCTGTGTGCTAGAGAGAAGGTATATTAGCAG	823
Qy	898	GAATATCTTCCGAGCTGCACGCTGCTGCTCTTTCAGATTGCCAAAAAGACCAAGAAAATG	957
Db	824	GTAATTCATCTGGTGTGCTGCATATATATGCTGCTATTCAAAAAGCAAAAAGATTAGGTAA--	881
Qy	958	CCGGGAAGCTTATTGTGTGCCGTTTTCCTCCAGCTTCGGGGAAGGTAACCTGTCCCTCCGTGC	1017
Db	882	-AGGTAAAAACAGTTGTAAACGATATTACCAAGTAATGGGAACGTTACTTATCAACAACCAT	940
Qy	1018	TATTT 1022	
Db	941	TATAT 945	

RESULT 9

```

US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

```

Query Match	Similarity	Score	DB 3	Length
Best Local	53.3%	Pred. No. 1,2e-52		
Matches	498	Conservative	0	Mismatches 433; Indels 3; Gaps 1;

  

Query	109	GAATTGCCAAGATGTTACGGAAATTGATGGTTAAACCCCATTTAGTATATCTTAATAAAC	168
Db	2604608	GCATGCCCGAGGACATCAACCACTCATCGGGCGCACACCGCTGTCGACTGCGCGGAG	2604667
Qy	169	TTGCCGATGGTTGTGTTGCCCGGGGTGCTCTAACTGGAGTTGATGGACCATGCTCTA	228
Db	2604668	TCACCGACGGCGCGTTGCCGACATCGTCCCAAGCTGGAATTTTCAACCGGCCAACA	2604727
Qy	229	GTGTGAAGGACAGGATTTGGGTATAGTATGATTTGCTGATGCAGAGGAAGGACTATCA	288
Db	2604728	GCGTAAAGACCGTATCGGGGTGCCATCTTCAAGCGCGGACGACGAGTTGATCA	2604787
Qy	289	CACCTGGAAGAGTGTCTCTCATTTGAGCCACAAGTGTATATCTGGCATTTGGATTAGCCT	348
Db	2604788	AGCCCGACACGATCATTTCTC--GAACCCACGAGCGGTAAACCGGCATCGCCCTGGCCA	2604844
Qy	349	TCATGGCAGACCCAGGGGTTACAAGTTCATTAATTACAATGCTGCTTCTATGAGTCTTG	408
Db	2604845	TGGTTTGGCGCGGCAACCGGCTACCGGTGCTGCTGACCATGCCCCGACGATGAGTCTGG	2604904

QY	409	AGAGAGAGATCATTTCTATTAGCTTTGGAGCTGAGTGGTTCTGACAGATCTCTGTAAG	468
Db	2604905	AGCGCCGGAATGTGCTCGCGGCATACGGTGTGTAATCTATCTGACTCCGGGTGCGGACG	2604964
QY	469	GAATGAAGGTGCTCTTCAGAAAGCTGAAGAGATATTGGCTAAGACGCCCAATGCCTACA	528
Db	2604965	GCATGTCAAGTGCCTATCGCCAAAGCTGAAGAGCTGGCCAAAGCCGATCAACGCTACTTCG	2605024
QY	529	TACTTCAACATTTTGAAAACCCCTGCCAATCCCCAAGTTTCATATTGAACCACTGCTCAG	588
Db	2605025	TGCCCCAGCAATTCGAGAACCCCGGCCAACCCTGCATTCATCGCGTCAACGACCGCGAGG	2605084
QY	589	AGATATGAAAAGGCTCCGATGGGAAAATTGATGCATTTGTCTTGGGATAGGCACGTGTG	648
Db	2605085	AGGTCTGGCGTGAACACCGACGGCAAGGTGACATCGTCTGCGGGAGTCCGCACCCGATG	2605144
QY	649	GTACAAATACAGGTGCTGTGAAAATATCTTAAAGAGCAGAATCCGATATATAAGCTGATTG	708
Db	2605145	GCACCAATCACCGCGCTCGCGCAGGTTCATCAAGGAACGCAAGCCGTCCGCCGTTCTGTG	2605204
QY	709	GTGTGAACCAAGTTGAAAGTCCAGTGTCTCAGAGGAAAGCTGTGCCACACAGATTTC	768
Db	2605205	CCGTAGAGCCGGCCGCGTCCGCGGTCTTTCTGTGGCCAGAAAGGACCGCACCCGATCC	2605264
QY	769	AAGGATGTGCTGTGTTTATCTCTGTGTCTTGAAGTCAATCTTGTGATGAAGTTG	828
Db	2605265	AGGCATCGCGCGCGGTTCTGTCGCCCGGTACTCGACAGAGCTTAGTCGACGAGATCA	2605324
QY	829	TTCAAAATACAGTGATGAAGCAATAGAACTGCMAAGCTTCTTCCGCTTAAAGAGGCC	888
Db	2605325	TTACCGTGGTAAAGAAAGCGCGCTCAACGTGGCGCGCGGCTGCCCCGGAAGAGGGCT	2605384
QY	889	TATTTGTGGGAATATCTTCCGAGCTGCAGCTGTCTGCTTTTCAGATTGCAAAAAGAC	948
Db	2605385	TGCTGTGGCATCTCTCCGGCGCGCGCACAGTGGCGCGCTTTCAGGTGGCCCCGCGCGC	2605444
QY	949	CAGAAAATGCCGGGAAGCTTATTGTTCGCGTTTTCACAGCTTCGGGAGAGGTAACCTGT	1008
Db	2605445	CAGAGAACCGCGGAAGCTAATCTGTGTAGTCTCCCCGCACTTCGGCGGAACGATATCTGA	2605504
QY	1009	CCTCCGTCTATTGAGTCACTGAACGCCGAAGC	1042
Db	2605505	GCACACCGTTGTTCGCCGACGTGCTGACTAAGC	2605538

## RESULT 10

```

; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLIBSCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire W.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

```

Query Match	16.5%;	Score 225.2;	DB 3;	length 4411529;
Best local Similarity	53.3%;	Pred. No. 1.2e-52;		
Matches 498;	Conservative 0;	Mismatches 433;	Indels 3;	Gaps 1
QY	109 GAATGCCAAGATGTTACGGAAATTGATTGTAATAACCCATTAGTATATCTAATTAAC	168		



Mon May 24 08:18:49 2004

us-09-931-457a-30.rtf

ප්‍රශ්න

Db	2608798	GCATCGCCCGAGCAATCACCCAACTCATCGGGCGCACACCCGCTGCTCCGACTGCGCCGAG	2608857
QY	169	TTGCGGATGTTGTGTGTCGCCGGGTGCTCTAACTGGAGTTGATGAGCCATGCTCTA	228
Db	2608858	TCACCGACGCGCCGCTTGCCGACATCGTCGCCAAGCTGGAATTTCTCAACCCGGCCAACA	2608917
QY	229	GTTGTGAAGGACAGCATTTGGGTATAGTATGATTTGCTGATGCAGAAAGAAAGGACTTATCA	288
Db	2608918	GCGTAAAGACCCTATTCGGGGTTGCCATGCTTCCAAAGCGGCCGAGCAGGAGGTTGATCA	2608977
QY	289	CACCTGAAAGAGTGTCTCATTTGAGCCCAACAAGTGTATATCTGGCATTTGATTAGCCT	348
Db	2608978	AGCCGACACGATCAATCTC--GAACCCAGAGCGGTAAACACCGGATTCGCCCTGGCCA	2609034
QY	349	TCATGCGACGACCGAGGGGTACAAAGCTCAATAATTACAATGCCCTGCTTCTATGAGTCTTG	408
Db	2609035	TGCTTTGCGCGGCAAGCGGCTACCGGTGCTGCTGAACCATGCCGACGATGAGTCTTG	2609094
QY	409	AGAGAAAGATCATTCTATTAGCTTTTGGAGCTGAGTTGGTTCTGCAGATCCTGCTAAG	468
Db	2609095	AGCGCCGATGTGCTGCGGCATACGGTGTGAATCATCTCTACCTCCGGGTGCGAGC	2609154
QY	469	GAATGAAGGTGCTGTTCAAGAAGCTGAAGAGATATTGGTAAAGCGCCCATGCTTACA	528
Db	2609155	GCATGTCAGGTGCCATTCGCCCAAGGCTGAGAGCTGGCCAAAGACGATCAAGCTACTTTCG	2609214
QY	529	TACTTCAACAATTTGAAAAACCTGCGCAATCCCAAGTTCATTATGAAAACACTGTCCAG	588
Db	2609215	TGCCCCAGCAATTGGAHAACCCGCGAACC CGCCATTCATCGCCTCAAGACCGCGAGG	2609274
QY	589	AGATATGGAAGGCTCCGATGGGAAATTGATGCAATTTGTTCTGGGATAGGCACTGGTG	648
Db	2609275	AGGTCTGGCGTGACACCGACGGCAAGGTGCACATCTCTGCGCGGAGTTCGGCACCGGTG	2609334
QY	649	GTAACAATAACAGTGTCTGGAAAAATATCTTAAAGACAGAAATCCGAATATAAAGCTGATTG	708
Db	2609335	GCACCAATCACCGCGCTCGCGCAGGTCAATCAAGGAACGAAAGCCGTCCGCCGCTTCTGTG	2609394
QY	709	GTTGTGAACCAAGTTGAAAGTCCAGTGTCTTCAGAGGAAAGCCTGTCCACACAAGATTG	768
Db	2609395	CCGTAGAGCCGCGCGCTGCGCTCTTTCTGTGGCCAGAAAGGAGCCGCAACCCGATCC	2609454
QY	769	AAGGATTTGCTGCTGTTTATATCCCTGTTGTTCTGCAAGTCAATCTTCTTGATGAAGTTG	828
Db	2609455	AGGCGATCGGCGCGGTTGTTCTCCGCGTACTGCAACAGAACTAGTGCAGAGATCA	2609514
QY	829	TTCAAAATATCAAGTATGAAGCAATAGAACTGCAAAAGCTTCTTGCGCTTAAAGAAAGGCC	888
Db	2609515	TTACCGTTCGTAACGAAGACGCTCAACGTGCGCGCCGCTGCGCCGGGAAGAGGCT	2609574
QY	889	TATTTGTGGGAATATCTTCCGAGCTGCAGCTGCTGCTTTTCAAGATTGCAAAAAGAC	948
Db	2609575	TGCTGTGCGGCATCTCTCGGGGCGCCGCCACAGTGGCGCTCTTCAAGTGGCCCGCGGC	2609634
QY	949	CAGAAATGCGGGAAGCTATTTGTTGCGCTTTTCCAGCTTCGGGAGAGTACCTGT	1008
Db	2609635	CAGAGAAACGCGGGAGAGCTAATGTCGTAAGTCTCCCGCACTTCGCGCAAGATATCTGA	2609694
QY	1009	CCTCCGTGCTATTTGAGTCAGTGAGACGCGAAGC	1042
Db	2609695	GCACACCGTTGTTGCGCCGACGTGCTGACTAAGC	2609728

## RESULT 1.1

US-09-557-884-1  
; Sequence 1, Application US/09557884

Patent No. 6506581  
GENERAL INFORMATION:

APPLICANT: Fleischmann et al.  
TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Human Genome Sciences, Inc.  
 STREET: 9410 Key West Avenue  
 CITY: Rockville  
 STATE: MD  
 COUNTRY: USA  
 ZIP: 20850  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3 1/2 inch diskette  
 COMPUTER: Dell Pentium  
 OPERATING SYSTEM: MS DOS V6.22  
 SOFTWARE: ASCII Text  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/557,884  
 FILING DATE: 25-Apr-2000  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/476,102  
 FILING DATE: JUN-5-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Michelle S. Marks  
 REGISTRATION NUMBER: 41,971  
 REFERENCE/DOCKET NUMBER: PB186P3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 301-309-8504  
 TELEFAX: 301-309-8439  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1830121 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-09-557-884-1

[illegible]

```

Db 1165688 CGCCGATTAATAGATTGTTGTAACAATCACTTCTGTTGCCGTTGAGCCAGTGA 1165747
Qy 726 AGTCC-----AGTCTCTCAGAGAGAAAGCTGCTCCACACAAAGTT 767
Db 1165748 TCTCCGTCATTAGTCAAACTTTAGCGGGTGAAGAGTAACCAAGCCCAACAAATT 1165807
Qy 768 CAAGGATGCTGCTGTTTATCCCTGCTGCTTGAAGTCAATCTTCTGATGAATT 827
Db 1165808 CAAGTATCGGTGCGGGTTTCATTCACAAAATTAGATTATCTATTATGATCGCGTA 1165867
Qy 828 GTTCAATATCAAGTATGAGCAATAGAACTGCAAAAGCTTCTGCGCTTAAGAAGGC 887
Db 1165868 GAAACTGTGATAGTATACCGCACTTCCACAGCTCGCTTAATGCGGAGAGAGGA 1165927
Qy 888 CTATTGTGGAATATCTTCCGAGCTGAGCTGCTGCTTTCAGATTGCAAAAAGA 947
Db 1165928 ATTCTGCAAGTATTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1165987
Qy 948 CCAGAAATGCGCGGAGCTTATTTGCTGCTTTCCTCCAGCTTCCGCGAGAGGTAACCTG 1007
Db 1165988 CCAGAAATGCTGATAACTCATTTGTTATTTGCTTCAAGCTCTGAACGCTACTTA 1166047
Qy 1008 TCCTCCGCTATTGA 1024
Db 1166048 AGCACAGCACTGTTGA 1166064

```

```

RESULT 12
US-09-643-990A-1
; Sequence 1, Application US/09643990A
; Patent No. 6528289
; GENERAL INFORMATION:
; APPLICANT: Robert D. Fleischmann
; Mark D. Adams
; Owen White
; Hamilton O. Smith
; J. Craig Venter
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; thereof, and Uses Thereof
;
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville,
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/643,990A
; FILING DATE: 23-Aug-2000
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,429
; FILING DATE: 1995-06-07
; APPLICATION NUMBER: 08/426,787
; FILING DATE: 1995-04-21
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB186P1C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-610-5790
; TELEFAX: 310-309-8439
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs

```

```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

```

```

Query Match          16.4%; Score 223.8; DB 4; Length 1830121;
Best Local Similarity 57.5%; Pred. No. 1.9e-52;
Matches 493; Conservative 0; Mismatches 337; Indels 27; Gaps 4;

```

```

Qy 192 GTTGCTCTAACTGAGTGTGATGAGCCATGCTCTAGTGTGAAGGACAGGATTGGTAT 251
Db 1165211 GTGGTGTAAATTTGAGGTGTGTAACCCAGCTACAGCGTAAATGCCGATTGGGCG 1165270
Qy 252 AGTATGATTGCTGATGAGAGAGAGAGGAGCTTATCACACCTGGAAGAGTGTCTCAT 311
Db 1165271 AATATGTGTGCAAGCAGAAAAAGATGCAAGCTCACAAAAGC--GAAAGAGTTGTA 1165327
Qy 312 GAGCCAAAGTGTGTAATCTGCAATTGATTAAGCTTCAATGCGAGCAGCGGCTTAC 371
Db 1165328 GATGCAACAGAGTGTGTAACGCGGCTTGTGCTTATGTTGCGGCGCTAGAGTTAT 1165387
Qy 372 AAGCTCATATTAACAAGCTGCTTCTATGAGTCTTGAAGAGATCATCTATTAGCT 431
Db 1165388 AAATCAAGCTCACTATGCAAGAAACATGATGTTGAAGAAAACGCTTATTGTCGGA 1165447
Qy 432 TTGAGAGCTGAGTGTGTTCTGACAGATCTGTAAGGGAATGAAGGTGCTGTCAAG 491
Db 1165448 TTGGGTGTAATTTAGTGTCTTACCGAAGGCGCAAAAGAGATGAAGGTGCTATTGCGAAA 1165507
Qy 492 GCTGAAGAGATATTGGCTAAGACGCCCATGCTA---CATACTTCAACAATTGAAAAC 548
Db 1165508 GCAGAAAGAAATTTGTTGCTTCTGATCCAGCGCTATGTCATGCTTAACAATTGAAAAT 1165567
Qy 549 CCTGCCAATCCCAAGGTTCTATTATGAACCACTGCTCCAGAGATATGGAAGGCTCCGAT 608
Db 1165568 CCAGCCCAACCCACAAATTTCTGAGAAACACAGGCGCTGAATTTGGAAGATACGAT 1165627
Qy 609 GGGAAATTTGATGATTTGTTCTGGGATAGGCACTGCTGTACATACAGG---TGCT 665
Db 1165628 GGCAGAGTCATGATTTGTTGTTGCTGCGGTAGGACAGAGTGTGATTAACGGCAATTTCT 1165687
Qy 666 GGAATATCTTAAAGACAGAAATCCGAATATTAAGCTGATTTGCTGGAACAGTTGA 725
Db 1165688 CGCGGATTAATTAAGTTTGTGTAACAATCACTTCTGTTGCCGTTGAGCCAGTGA 1165747
Qy 726 AGTCC-----AGTCTCTCAGAGAGAAAGCTGTGTCACACAAAGATT 767
Db 1165748 TCTCCGTCATTAGTCAAACTTTAGCGGGTGAAGAGTAACCAAGGCCACACAAATTT 1165807
Qy 768 CAAGGATGCTGCTGCTGTTTATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 827
Db 1165808 CAAGTATCGGTGCGGGTTTCATTCACAAAATTAGATTATCTATTATGATCGCGTA 1165867
Qy 828 GTTCAATATCAAGTATGAGCAATAGAACTGCAAAAGCTTCTGCGCTTAAGAAGGC 887
Db 1165868 GAAACTGTGATAGTATACCGCACTTCCACAGCTCGCTTAATGCGGAGAGAGGA 1165927
Qy 888 CTATTGTGGAATATCTTCCGAGCTGAGCTGCTGCTGCTTTCAGATTGCAAAAAGA 947
Db 1165928 ATTCTGCAAGTATTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1165987
Qy 948 CCAGAAATGCGCGGAGCTTATTTGCTGCTTTCCTCCAGCTTCCGCGAGAGGTAACCTG 1007
Db 1165988 CCAGAAATGCTGATAACTCATTTGTTATTTTGCCTTCAAGCGTCTGAACGCTACTTA 1166047
Qy 1008 TCCTCCGCTATTGA 1024
Db 1166048 AGCACAGCACTGTTGA 1166064

```

```

RESULT 13
US-09-790-988-1/c

```

Sequence 1, Application US/09790988  
Patent No. 6632935  
GENERAL INFORMATION:  
APPLICANT: SHIGENOBU, SHUJI  
APPLICANT: MATANABE, HIDEMI  
APPLICANT: HATTORI, MASAHIRA  
APPLICANT: SAKAKI, YOSHIYUKI  
TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS  
FILE REFERENCE: 081356/0159  
CURRENT APPLICATION NUMBER: US/09/790,988  
CURRENT FILING DATE: 2001-02-23  
PRIOR APPLICATION NUMBER: JP2000-107160  
PRIOR FILING DATE: 2000-04-07  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 640681  
TYPE: DNA  
ORGANISM: Buchnera sp.  
US-09-790-988-1

Query Match 16.1%; Score 218.8; DB 4; Length 640681;  
Best Local Similarity 56.2%; Pred. No. 2.9e-51;  
Matches 534; Conservative 0; Mismatches 377; Indels 39; Gaps 5;

QY 100 AAGGTCGGAATGTCGAAGATGTTACGAATTGATGTTAAACCCCATTAATATATC 159  
DB 72179 AATGAGTAAATATATATCAAGATAATCTTAACTATGTAATACACCTCTGTCCTT 72120  
QY 160 TAAATAAATTGCGGATGTTGTTGCGCGGTTGCTGCTAACTGAGTTGATGAGC 219  
DB 72119 TAAATAGATTTGGAATGGAAT-----ATTAGTAAATAAGATCTAGAAATC 72069  
QY 220 CATGCTCTAGTGTGAAGACAGATGGGTATAGTATGATGCTGATGAGAGAGAG 279  
DB 72068 CAGCTTTAGTGTAAATGTAATGTCCTAATATGATATGACGCGGAGAGAGAG 72009  
QY 280 GACTTATCACACCTGGAAGAGTCTCTCATTTGAGCCCAAGTGTATATCTGGCATG 339  
DB 72008 AATAATATAA---TAAATATGTAATACTAATCGAGGCGACTAGTGTATATCAGAGATAG 71952  
QY 340 GATTAAGCTTTCATGCGAGCAGCGGCTTACAGCTCATTAATCAATGCTGCTCTA 399  
DB 71951 CATTAAGCTTATGTTGCGAGCTTCTGAAATATATGATTAATCTAATGCTGAAACCA 71892  
QY 400 TGAGTCTTGAGAGAGATCTTCTATTAAGCTTTTGGAGCTGAGTTGTTCTGAGATC 459  
DB 71891 TGCTATTTGAGCGGAGAGAGAGATTAATCTTTAGGTCGAGATTTGATTTTACCGATG 71832  
QY 460 CTGCTAAGGGAATGAAGGTCCTGTTGAGAGGCTGAGAGATATTGGCTAAGAGCCCA 519  
DB 71831 GAAGATATGCGATGAGAGGAGCTATTCTAAGGCAATGATATATATCTGCAATCCAA 71772  
QY 520 ATGCTA---CATACTTCAACAATTGAAAAACCTGCGCAATCCCAAGTTCTATATGAAA 576  
DB 71771 GTAAATATTTTATTATTAACAATTGAGAAATCCAGCTAATCTGAAATTCATCAAAATTA 71712  
QY 577 CCACTGCTCCAGAGATATGAGAGGCTCCGATGGGAAATGATGCAATTTGTTCTGGGA 636  
DB 71711 CTACTGACCAAGAAATATGAGACATTAATGCAATTTAGATATTATTTATTTCTGCTG 71652  
QY 637 TAGGCACTGCTGTACATAAAGCTGCTGAAAAATATCTTAAAGAGCAGAAATCCGATA 696  
DB 71651 TAGGAACGGGTGCACAATTAATGTAATCAAGATATATCAAAAAATTAAGAAAAA 71592  
QY 697 TAAAGCTGATGTTGTTG-----GAACCACTTGAAGAGTCCAGTCTCTCAG 741  
DB 71591 AAAATTTAATCAGTATAGCTGTGAACCTTCTGATCACTGTAATCAACAATTTTAG 71532  
QY 742 GAGGAA-----AGCTGCTCAACAAGATTCAGAGGATGCTGCTGTTTATCC 792  
DB 71531 CAGGAAAGCAATCGAGCTGAGCAATTAATTAAGAGATGGGCTGTTTATATAC 71472

QY 793 CTGCTGTTGGAAGTCAATCTTCTGATGAGAGTGTTCATAATATCAAGTATGAGCAA 852  
DB 71471 CCAAAACTTAGATTTAACAATATATGATCAAGTATATCTGTTCTAGTGAAGAGCAA 71412  
QY 853 TAGAACTGCAAAAGCTTCTTGGCTTAAAGAGGCTATTGTTGGAATATCTCCGAG 912  
DB 71411 TATTACGGCTAAAGATTAATGAAAAAGAGGATATAGCAGAGAACTCTCTGCTG 71352  
QY 913 CTGACGCTGCTGCTCTTTTCAATGCAAAAAAGACCAAGAAATGCGGGAAGCTTATG 972  
DB 71351 CTGCGTTATATGCAATTAATAATCCACAGCAAAAAATTTTCAACAAAAATAG 71292  
QY 973 TTGCCGTTTCCAGCTTTCGGGAGAGGTAACCTGCTCCGTTGCTATT 1022  
DB 71291 TTGTTATTTAACCCTTCTCAGAGAGAGGTTATTAGTACAGAAATTAAT 71242

RESULT 14  
US-09-724-623-10  
Sequence 10, Application US/09724623  
Patent No. 6476209  
GENERAL INFORMATION:

APPLICANT: Glenn, Matthew  
APPLICANT: Lubbers, Mark W  
APPLICANT: Dekker, James  
TITLE OF INVENTION: Polynucleotides, materials incorporating  
FILE REFERENCE: 104801  
CURRENT APPLICATION NUMBER: US/09/724,623  
CURRENT FILING DATE: 2000-11-28  
NUMBER OF SEQ ID NOS: 124  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10  
LENGTH: 927  
TYPE: DNA  
ORGANISM: Lactobacillus rhamnosus  
US-09-724-623-10

Query Match 15.4%; Score 209.8; DB 4; Length 927;  
Best Local Similarity 55.3%; Pred. No. 3.1e-50;  
Matches 472; Conservative 0; Mismatches 372; Indels 9; Gaps 3;

QY 114 GCCAAGATGTTACGAATGATTTGTTAAACCCCATTAATATCTAATAAATGCG 173  
DB 13 GCAGATATATTAACAGGTTAATTTGCAATACCGCTACTCAAGCTCAATCGCTTGA 72  
QY 174 GATGTTGTTGTCGCGGCTGCTGCTAACTGAGTGTATGAGAGCCATGCTAGTGT 233  
DB 73 CCGAAGGCGCGCGGATGTTTATGTCAGCTGGAATCTTTAATCCGCGGCTCAGTC 132  
QY 234 AAGACAGATGGGTATATGATGATTTGCTGATGCAAGAGAGAGAGGACTTATCACACT 293  
DB 133 AAGACCGGATGCTTGGCGATGATGAGAGCGCTGAATATTAAGGGCTTGAAGCCA 192  
QY 294 GGAAGAGTGTCTCATTTAGGCCAACAGTGTATATCTGCAATTTGATGCTTATG 353  
DB 193 GG---CGGACCAATGTTGAGGCCAACGTCGCGCAACACCGGCAATTTGAGCTG 249  
QY 354 GCAGACGCGAGGTTACAGCTCAATTAATCAATGCTGCTTCTATGAGCTTTGAGAGA 413  
DB 250 GCGGCGGCAAAAGTTATCACTCATCAATCAATGCGGAAACGATGAGTGTGAGCGG 309  
QY 414 AGAATCATTTATAGCTTTTGGAGCTGAGTGTGCTGACAGATCTGCTAAGGGAATG 473  
DB 310 CGTCTTTGATGCGTGTGTACGAGCGCACTCAATTTGACGCGGCTGCGATGGAATG 369  
QY 474 AAGGTGCTGTTCAAGAGGCTGAGAGATATTGCTAAGAGCGCCCAATGCTTACATCT 533  
DB 370 CCGGAGCAATTAATAAAGCAGAG---CATTAAGCAGAGAAATGCTACTTTCGCA 426  
QY 534 CAACAATTTGAAAACTGCGCAATCCCAAGGTTCAATATGAAACCACTGCTCAGA---G 590  
DB 427 ATGCAATTTCAAGAACCCCGCAATCCAGAGCTCCACGAGCGCACGACGAGCAAGAAATC 486



QY	591	ATATGGAAGGCTCCCGATGGGAAAATTGATGCATTTGTTCTTGGGATAGCACTGGTGT	650
Db	487	ATCCGTTCAATTGATGTGTGGCACCCGAGATGCCCTTTGTAGCCGGCGGTGGGACAGGCGGA	546
QY	651	ACAATAACAGGTGCTGGAAAATATCTTAAAGACGACAGATCCGATATTAAGCTGATTGGT	710
Db	547	ACACTCACCGGGTTGGTGGGCTCTGGCTAAGATCAATCCAGATGTACAAATCTATGGG	606
QY	711	GTGAAACCAGTTGAAAGTCCAGTGTCTTCAGGAGGAAAAGCCTGGTCCACACAGATTCAA	770
Db	607	TTCGAAACGACGGAGTGCCTCAATGCTTAAAGAGGCCCATGGCGGCAAGCACAGATTCAA	666
QY	771	GGGATTGTGCTGGTTTATCCCTGGTGTCTTGGAAAGTCAATCTTCTTGATGAAGTTGTT	830
Db	667	GGGATCTCAGCCGGTTTATTCAGACGCTTTAGATACGAACTCTATCAAGCATCATTT	726
QY	831	CAATATCAAGTGATGAAGCAATAGAAACTGCAAAAGCTTCTTGGCTTAAAGAAAGCCTTA	890
Db	727	GAAATCACACAGCGATCAAGCTATGACATGGCTGGCCACGTCAGCCATGAAGAGGCTTC	786
QY	891	TTTGGGGAATATCTTCGGAGCTGACGTGCTGTCTTTTCAGATTGCAAAAAGACCA	950
Db	787	CTACCAAGGCAATTCCGCTGGCGCTAACATTTTGGCGCGATTGAANTCGCCAAAGAACTC	846
QY	951	GAATAATGCCGGGA	963
Db	847	GGCAAAGGCAAGA	859

RESULT 15  
US-09-634-238-26/c

; Sequence 26, Application US/09634238  
; Patent No. 6544772

GENERAL INFORMATION:

APPLICANT: Glenn, Matthew  
ADDRESS: 1701

APPLICANT: HAVUKKALA, IIRKA O.  
APPLICANT: BLOKSBERG, LEONARD N.

APPLICANT: Rubbers, Mark W.

APPLICANT: Dekker, James

APPLICANT: Christensson, Anna C.

APPLICANT: Holland, Ross

APPLICANT: O'Toole, Paul W.

APPLICANT: Reid, Juelan K.  
APPLICANT: Timothybear

TITLE OF INVENTION: polynucleotides, materials incorporating

TITLE OF INVENTION: them and methods for using them.

FILE REFERENCE: 11000.1043U1

CURRENT APPLICATION NUMBER: US/00

CURRENT FILING DATE: 2000-08-08  
NUMBER OF SEC ID NOS: 433

NUMBER OF SEQ ID NOS: 422  
SOFTWARE: FastSeq for Windows Version 4.0

DOF IMAGE: F8  
 SED ID NO 26

LENGTH: 8885

TYPE: DNA

ORGANISM: *Lactobacillus rhamnosus*;

US-09-634-238-26

Query Match	15.4%	Score 209.8;	DB 4;	Length 8885;
Best Local Similarity	55.3%;	Pred. No. 1.1e-49;		
Matches 472;	Conservative	0;	Mismatches 372;	Indels 9;
				Gaps 3;

OY	114	GCCAAAGATGTTACGGAAATTGATGGTAAAAAACCOCATTAGTATATCTTAATAACTTGCG	173
Dp	5600	GCA GATAATATTACAGGTTTAATTGCACAFTACGCCGCTACTCAAGCTCAATCGCGTTGTA	5541
OY	174	GATGCTGTGTGTCGCCGGTGTCCTCTAACTGGAGTTGATGAAGCCATGCTCTAGTGTG	233
Dp	5540	CCTGAAGCGCGCCGGATGTTATATCACAAGCTGGAATCTTTAATCCGGCGGCTCAGTC	5481
OY	234	AAGCACAGATTTGGGTATAGTATGATTCGTGATGCAGAAGAGAGGAACTTATCACACT	293
Dp	5480	AAGAACCGGATTTGCCCTTGCGCATGATTGAAGACGCTGAATATATAAGGGGCTTGAAGCCA	5421

QY	294	GGAAAGAGTGTCTCTCATTTGAGCCACAAAGTGTATATACTGGCATTGGATTAGCCCTCATG	353
Db	5420	GG--CGGCACCATTGTGTGAGCCAAACGTCGCCGACAACCGGCATTGGACTGGCACTGGTT	5364
QY	354	GCAGCAGCCAGGGGTTACAAGCTCATATATTACATATGCTGCTCTTATGAGTCTTGAAGCA	413
Db	5363	GCGCGGCGAAAGTTATTCACCTCATCATCACCATGCGGAAACGATGAGTGTGAGCGG	5304
QY	414	AGAATCATTCTATTAGCTTTTGGAGCTGAGTTGGTTCTGACAGATCTGCTTAAGGAATG	473
Db	5303	CGTGCTTTGATGCGTGTGTACGGAGCCGAACCTCATTTTGAACGCCGGGTGCCGATGGAATG	5244
QY	474	AAAGTGCTGTTCAGAAAGCTGAAGAGATATTGGCTAAGACGCCCAATGCTTACATACTT	533
Db	5243	CCGGAGCATTTAAAAAGACAAG--CATTAAGCAAGAAATGGTACTACTTCTTGCCA	5187
QY	534	CAACATTTGAAAAACCTGCCAATCCCAAGTTCATTATGAACCACTGTCAGA---G	590
Db	5186	ATGCAATTCAGAACCCCGCCCAATCCAGACGTCCACGACGCCACGACCGACAGAAATC	5127
QY	591	ATATGAAAAGGCTCCGATGGGAAATATGATGCATTGTGTTCTGGGATAGGCACTGGTGT	650
Db	5126	ATCCGTTCAATTGATGTGTGACCCCGAGATGCTTTGTAGCCGGCGCTGGCACAGCGGA	5067
QY	651	ACAATAACAGGTGCTGGAATAATCTTAAGAGCAGATCCGATATTAAGCTGATTGT	710
Db	5066	AACTCAACCGGGTGTGGGCTCTGCGTAAGATCAATCCAGATGTACAATCTTATGCG	5007
QY	711	GTGGAACCAAGTTGAAAGTCCAGTGTCTCAGAGAGAAAGCTGTCCACACAAAGATTCAA	770
Db	5006	TTGGAAGCAGCGGAGTCCGCCAATGCTTAAAGAAAGGCCATGGCGCAAGCACAAGATTCAA	4947
QY	771	GGGATTGCTGTGTTTATCCCTGCTGTCTTGAAGTCAATCTTCTTGATGAAGTTGT	830
Db	4946	GGGATCTCAGCCGGTTTATTCAGACGCTTTAGATTACGAACCTCTATCAAGACATCAT	4887
QY	831	CAATATCAAGTGAAGCAATGAAACTGCAAAGCTTCTTGCGCTTAAGAAGGCCCTA	890
Db	4886	CAAGTCAACCGCATCAAGCTATCGACATGGCTGGCCACGTCAGCCATGAAGAAGGCTTC	4827
QY	891	TTTGTGGAATATCTTCCGAGCTGCAGCTGCTGCTTTTCAGATTGCCAAAAAGACCA	950
Db	4826	CTAACGAGCATTTCCGCTGGCGCTAAACATTTTGGCGGATTTGAATCGCCAAAGAACTC	4767
QY	951	GAATATGCCGGA	963
Db	4766	GGCAAAAGCAAGA	4754

Search completed: May 21, 2004, 20:35:46  
Job time : 159 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: May 21, 2004, 19:00:02 ; Search time 703 Seconds  
(without alignments)  
8805.006 Million cell updates/sec

Title: US-09-931-457A-30  
Perfect score: 1362  
Sequence: 1 actttgtagtcgtagatag.....aaaaaaaaaaaaa 1362

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2953838 seqs, 2272363821 residues

Total number of hits satisfying chosen parameters: 5907676

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

1:	/cgn2_6/ptoddata/2/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/ptoddata/2/pubpna/PCT_NEW_PUB.seq:*
3:	/cgn2_6/ptoddata/2/pubpna/US06_NEW_PUB.seq:*
4:	/cgn2_6/ptoddata/2/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/ptoddata/2/pubpna/US07_NEW_PUB.seq:*
6:	/cgn2_6/ptoddata/2/pubpna/PCTUS_PUBCOMB.seq:*
7:	/cgn2_6/ptoddata/2/pubpna/US08_NEW_PUB.seq:*
8:	/cgn2_6/ptoddata/2/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/ptoddata/2/pubpna/US09A_PUBCOMB.seq:*
10:	/cgn2_6/ptoddata/2/pubpna/US09B_PUBCOMB.seq:*
11:	/cgn2_6/ptoddata/2/pubpna/US09C_PUBCOMB.seq:*
12:	/cgn2_6/ptoddata/2/pubpna/US09_NEW_PUB.seq:*
13:	/cgn2_6/ptoddata/2/pubpna/US09_NEW_PUB.seq2:*
14:	/cgn2_6/ptoddata/2/pubpna/US10A_PUBCOMB.seq:*
15:	/cgn2_6/ptoddata/2/pubpna/US10B_PUBCOMB.seq:*
16:	/cgn2_6/ptoddata/2/pubpna/US10C_PUBCOMB.seq:*
17:	/cgn2_6/ptoddata/2/pubpna/US10_NEW_PUB.seq:*
18:	/cgn2_6/ptoddata/2/pubpna/US60_NEW_PUB.seq:*
19:	/cgn2_6/ptoddata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1362	100.0	1362	9	US-09-931-457A-30	Sequence 30, Appl
2	1340.6	98.4	1407	13	US-10-424-599-130460	Sequence 130460, A
3	660.8	48.5	3075	13	US-10-424-599-55164	Sequence 55164, A
4	595.2	43.7	966	9	US-09-938-842A-893	Sequence 893, App
5	595.2	43.7	966	11	US-09-938-842A-893	Sequence 893, App
6	578.6	42.5	1404	13	US-10-424-599-55160	Sequence 55160, A
7	557.4	40.9	1319	13	US-10-425-114-25435	Sequence 25435, A
8	555.8	40.8	1308	13	US-10-425-114-19467	Sequence 19467, A
9	555.8	40.8	1312	13	US-10-425-114-23298	Sequence 23298, A
10	555.8	40.8	1403	13	US-10-425-114-1445	Sequence 1445, Ap
11	554.2	40.7	1301	13	US-10-425-114-23608	Sequence 23608, A
12	543.4	39.9	1603	13	US-10-424-599-61234	Sequence 61234, A
13	524.4	38.5	1279	13	US-10-425-114-22525	Sequence 22525, A
14	521.8	38.3	1490	13	US-10-425-114-34372	Sequence 34372, A

15	521.8	38.3	1491	13	US-10-425-114-34527	Sequence 34527, A
16	521.6	38.3	1522	9	US-09-931-457A-63	Sequence 63, Appl
17	520.2	38.2	1508	9	US-09-931-457A-61	Sequence 61, Appl
18	520.2	38.2	1509	13	US-10-425-114-30493	Sequence 30493, A
19	495.8	36.4	975	10	US-09-905-290A-3	Sequence 3, Appl1
20	477.2	35.0	972	9	US-09-938-842A-495	Sequence 495, App
21	477.2	35.0	972	11	US-09-938-842A-495	Sequence 495, App
22	477.2	35.0	1472	9	US-09-887-576-611	Sequence 611, App
23	476.4	35.0	1329	9	US-09-938-842A-99	Sequence 99, Appl
24	476.4	35.0	1329	11	US-09-938-842A-99	Sequence 99, Appl
25	434.8	31.9	2307	13	US-10-425-114-7548	Sequence 7548, Ap
26	427.8	31.4	1925	13	US-10-424-599-18037	Sequence 18037, A
27	396.8	29.1	1563	13	US-10-424-599-53377	Sequence 53377, A
28	390.8	28.7	594	13	US-10-424-599-47440	Sequence 47440, A
29	385.4	28.3	588	16	US-10-341-961A-271	Sequence 271, App
30	353.6	26.0	808	13	US-10-424-599-18732	Sequence 18732, A
31	335.2	24.6	528	13	US-09-770-152-879	Sequence 879, App
32	329.6	24.2	1095	13	US-10-424-599-99835	Sequence 99835, A
33	327.2	24.0	930	13	US-10-282-122A-15848	Sequence 15848, A
34	322.6	23.7	948	13	US-10-424-599-103489	Sequence 103489, A
35	308	22.6	962	9	US-09-770-445-305	Sequence 305, App
36	303.2	22.3	927	13	US-10-282-122A-16019	Sequence 16019, A
37	300.4	22.1	542	13	US-09-770-152-652	Sequence 652, App
38	294	21.6	1021	13	US-10-424-599-109938	Sequence 109938, A
39	280.8	20.6	945	13	US-10-282-122A-12236	Sequence 12236, A
40	276.2	20.3	568	9	US-09-734-017A-27	Sequence 27, Appl
41	273	20.0	927	13	US-10-282-122A-26908	Sequence 26908, A
42	273	20.0	100848	13	US-10-672-787-39	Sequence 39, Appl
43	262.6	19.3	529	13	US-10-424-599-55162	Sequence 55162, A
44	258.6	19.0	373	13	US-10-424-599-127232	Sequence 127232, A
45	256.4	18.8	930	13	US-10-282-122A-21035	Sequence 21035, A

## ALIGNMENTS

RESULT 1  
US-09-931-457A-30  
Sequence 30, Application US/09931457A  
Patent No. US20020157132A1  
GENERAL INFORMATION:  
APPLICANT: Falco, S. Carl  
TITLE OF INVENTION: Plant Amino Acid Biosynthetic Enzymes  
FILE REFERENCE: B1116 US CIP  
CURRENT APPLICATION NUMBER: US/09/931,457A  
CURRENT FILING DATE: 2002-02-22  
PRIOR APPLICATION NUMBER: 09/424,976  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: 60/065,385  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/049,406  
PRIOR FILING DATE: 1997-06-12  
NUMBER OF SEQ ID NOS: 72  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 30  
LENGTH: 1362  
TYPE: DNA  
ORGANISM: Glycine max  
US-09-931-457A-30

Query Match 100.0%; Score 1362; DB 9; Length 1362;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ACTTTGTAGTTCGTAGATAGCCGATGCTGCTTCTTAGTGTGTCAGTCAATCCTGTTCT	60
DB	1	ACTTTGTAGTTCGTAGATAGCCGATGCTGCTTCTTAGTGTGTCAGTCAATCCTGTTCT	60
QY	61	CAAGTCAAGCTTTAGTGAAGCAGATATATGCTGTTGAAGGTCGGAATTGCCAAG	120
DB	61	CAAGTCAAGCTTTAGTGAAGCAGATATATGCTGTTGAAGGTCGGAATTGCCAAG	120

QY 121 ATGTTACGGAATGATGGTAAACCCCATTTAGTATATCTAAATAAAGTTGGGATGTT 180  
 DB 121 ATGTTACGGAATGATGGTAAACCCCATTTAGTATATCTAAATAAAGTTGGGATGTT 180  
 QY 181 GTGTGCCCCGGGTGCTGCTAACTGAGGTGATGAGCCATGCTCTAGTGTGAAGACA 240  
 DB 181 GTGTGCCCCGGGTGCTGCTAACTGAGGTGATGAGCCATGCTCTAGTGTGAAGACA 240  
 QY 241 GGATGGGTATAGTATGATGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
 DB 241 GGATGGGTATAGTATGATGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
 QY 301 GTGTCTCATTTAGGACCAACAGGTGTAATGCTGCAATGGAATGAGCTTCAATGAGCAG 360  
 DB 301 GTGTCTCATTTAGGACCAACAGGTGTAATGCTGCAATGGAATGAGCTTCAATGAGCAG 360  
 QY 361 CCAGGGGTACAGCTCATATAATCAATGCTGCTCTAGTGTGAAGAGAGAGATCA 420  
 DB 361 CCAGGGGTACAGCTCATATAATCAATGCTGCTCTAGTGTGAAGAGAGAGATCA 420  
 QY 421 TTCTATTAGCTTTTGGAGCTGAGTGGTTCTGACAGATCCTGCTAAGGAGAGAGAG 480  
 DB 421 TTCTATTAGCTTTTGGAGCTGAGTGGTTCTGACAGATCCTGCTAAGGAGAGAGAG 480  
 QY 481 CTGTTTCAAGAGGCTGAGAGAGATTTGGCTAAGACGCCCAATGCTTCAATCAACAT 540  
 DB 481 CTGTTTCAAGAGGCTGAGAGAGATTTGGCTAAGACGCCCAATGCTTCAATCAACAT 540  
 QY 541 TTGAAAAACCTGCGCAATCCCAAGTTCATTTATGAACCACTGCTCCAGAGATGAGAA 600  
 DB 541 TTGAAAAACCTGCGCAATCCCAAGTTCATTTATGAACCACTGCTCCAGAGATGAGAA 600  
 QY 601 GCTCCGATGGAAAAATGATGATGCTTTGTTCTGGAGTAGGCACTGTTGTTACATTAAC 660  
 DB 601 GCTCCGATGGAAAAATGATGATGCTTTGTTCTGGAGTAGGCACTGTTGTTACATTAAC 660  
 QY 661 GTGCTGGAATAATATCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
 DB 661 GTGCTGGAATAATATCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
 QY 721 TTGAAAGTCCAGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
 DB 721 TTGAAAGTCCAGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
 QY 781 CTGCTTTTATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
 DB 781 CTGCTTTTATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
 QY 841 GTGATGAAGCAATAGAAAGTGAAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
 DB 841 GTGATGAAGCAATAGAAAGTGAAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
 QY 901 TATCTTCCGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
 DB 901 TATCTTCCGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
 QY 961 GGAAGCTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
 DB 961 GGAAGCTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
 QY 1021 TTGAGTCAAGTGAAG 1080  
 DB 1021 TTGAGTCAAGTGAAG 1080  
 QY 1081 GCTCTCACTACTGAAATTTCTGTTAATCTGTAACAGGCTTTAATCTGTAAGAGTAC 1140  
 DB 1081 GCTCTCACTACTGAAATTTCTGTTAATCTGTAACAGGCTTTAATCTGTAAGAGTAC 1140  
 QY 1141 TACTGTTTGTGACTCTGACTCTAAATAAAGTGTGCTCAAAAGTGTGTTTCTGAT 1200  
 DB 1141 TACTGTTTGTGACTCTGACTCTAAATAAAGTGTGCTCAAAAGTGTGTTTCTGAT 1200  
 QY 1201 GCCCCTGAGAGGATATTTTGTGCTGCAACATTAATAAGTATTCAAAGTGTGTTAAG 1260

DB 1201 GCCCCTGAGAGGATATTTTGTGCTGCAACATTAATAAGTATTCAAAGTGTGTTAAG 1260  
 QY 1261 TAACATGTTTCACTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1320  
 DB 1261 TAACATGTTTCACTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1320  
 QY 1321 TTCTGATTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1362  
 DB 1321 TTCTGATTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1362

RESULT 2  
 US-10-424-599-130460  
 ; Sequence 130460, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa Thomas J  
 ; APPLICANT: Kovalic David K  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 265684  
 ; SEQ ID NO 130460  
 ; LENGTH: 1407  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_88813C.1  
 ; US-10-424-599-130460

Query Match 98.4%; Score 1340.6; DB 13; Length 1407;  
 Best Local Similarity 99.3%; Pred. No. 0;  
 Matches 1346; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
 QY 2 CTTGTAGTCTGATAG 61  
 DB 38 CTTGTAGTCTGATAG 97  
 QY 62 AAGTCAAGCTTTGATGAG 121  
 DB 98 AAGTCAAGCTTTGATGAG 157  
 QY 122 TGTAGGAATGATGTTGTTAAACCCCATTAATATCTAATAAATCTGCGATGTTG 181  
 DB 158 TGTAGGAATGATGTTGTTAAACCCCATTAATATCTAATAAATCTGCGATGTTG 217  
 QY 182 TGTGCCCCGGGTGCTGCTAACTGAGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAG 241  
 DB 218 TGTGCCCCGGGTGCTGCTAACTGAGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAG 277  
 QY 242 GATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 301  
 DB 278 GATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 337  
 QY 302 TGTCTCATTTAGGACCAACAGGTGTAATGCTGCAATGGAATGAGCTTCAATGAGCAGC 361  
 DB 338 TGTCTCATTTAGGACCAACAGGTGTAATGCTGCAATGGAATGAGCTTCAATGAGCAGC 397  
 QY 362 CAGGGGTACAGCTCATATAATCAATGCTGCTTCTATGAGTGTGAGAGAGAGATCAT 421  
 DB 398 CAGGGGTACAGCTCATATAATCAATGCTGCTTCTATGAGTGTGAGAGAGAGATCAT 457  
 QY 422 TCTATTAGCTTTTGGAGCTGAGTGGTTCTGACAGATCCTGCTAAGGAGATGAAGGTGC 481  
 DB 458 TCTATTAGCTTTTGGAGCTGAGTGGTTCTGACAGATCCTGCTAAGGAGATGAAGGTGC 517  
 QY 482 TGTTCAGAGGCTGAAGAGATATTTGCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAT 541



Db 518 TGTTGAGAGGCTGAGAGATATTGGCTAAGACGCCCAATGCTTACATCTTCAACATT 577  
QY 542 TGAAGACCTGCTCCATCCCAAGGTTCTATTATGAACCACTGCTCCAGAGATATGAAAG 601  
Db 578 TGAAGACCTGCTCCATCCCAAGGTTCTATTATGAACCACTGCTCCAGAGATATGAAAG 637  
QY 602 CTCCGATGGGAAATGATGATTTGTTTCTGGGATAGGCACTGGTGTGATCAATACAG 661  
Db 638 CTCCGATGGGAAATGATGATTTGTTTCTGGGATAGGCACTGGTGTGATCAATACAG 697  
QY 662 TGCTGAAATATCTTAAAGAGCAGATCCGATATTAAGCTGATTGGTGGAAACAGT 721  
Db 698 TGCTGAAATATCTTAAAGAGCAGATCCGATATTAAGCTGATTGGTGGAAACAGT 757  
QY 722 TGAAGTCCAGTCTCTCAGAGAGAAAGCTGTGTCACACAGATTCAGGATTTGTC 781  
Db 758 TGAAGTCCAGTCTCTCAGAGAGAAAGCTGTGTCACACAGATTCAGGATTTGTC 817  
QY 782 TGGTTTATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 841  
Db 818 TGGTTTATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 877  
QY 842 TGATGAAGCAATAGAACTGCAAGCTTCTGCGCTTAAAGAGGCTATTGTGGAAAT 901  
Db 878 TGATGAAGCAATAGAACTGCAAGCTTCTGCGCTTAAAGAGGCTATTGTGGAAAT 937  
QY 902 ATCTCCGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 961  
Db 938 ATCTCCGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 997  
QY 962 GAAGCTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1021  
Db 998 GAAGCTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1057  
QY 1022 TGAGTCAAGTACGCGAGAGCTGAAGAGTACTTTTGAAGCTTCCGTTAAG 1081  
Db 1058 TGAGTCAAGTACGCGAGAGCTGAAGAGTACTTTTGAAGCTTCCGTTAAG 1117  
QY 1082 CTCTCACTACTGAAATTTCTGTTACTGTTACCAAGCTTTAATGTTAGAGTACT 1141  
Db 1118 CTCTCACTACTGAAATTTCTGTTACTGTTACCAAGCTTTAATGTTAGAGTACT 1177  
QY 1142 ACTGTTTGAATCTGACTCTAAATATAAATCTGCTCCAAAGAGTACTTTTCTGATG 1201  
Db 1178 ACTGTTTGAATCTGACTCTAAATATAAATCTGCTCCAAAGAGTACTTTTCTGATG 1237  
QY 1202 CCCCCTGAGCGATTAATTTTGTGCTGCAACATTAAGATTAAGTCTTATAGT 1261  
Db 1238 CCCCCTGAGCGATTAATTTTGTGCTGCAACATTAAGATTAAGTCTTATAGT 1297  
QY 1262 AACATGTTCACTTTTGTGTTGTTGAGACGAACAGGATGAGTCAATATAGT 1321  
Db 1298 AACATGTTCACTTTTGTGTTGTTGAGACGAACAGGATGAGTCAATATAGT 1357  
QY 1322 TCTGATTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1356  
Db 1358 TCTGATTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1392

RESULT 3  
US-10-424-599-55164  
; Sequence 55164, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223) B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 55164  
; LENGTH: 3075  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURES:  
; OTHER INFORMATION: Clone ID: PAT\_MFT3847\_20824C.1  
US-10-424-599-55164

Query Match 48.5%; Score 660.8; DB 13; Length 3075;  
Best Local Similarity 77.8%; Pred. No. 6.4e-176;  
Matches 797; Conservative 0; Mismatches 227; Indels 0; Gaps 0;

QY 86 TATATGCTGTTGAAGTCCGAAATTCGCAAGATGTTACGGAATTTGATTTGTAAC 145  
Db 308 TATATGCTGTTGAAGTCCGAAATTCGCAAGATGTTACGGAATTTGATTTGTAAC 367  
QY 146 CCAATTAATATCTAATATAAATTCGCAAGATGTTGTTGTTGTTGTTGTTGTTGTTGTT 205  
Db 368 TCAATGTTATATCTCAACCGCATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 427  
QY 206 GAGTTGATGAGCCATGCTCTAGTGTGAAGACAGAGATGGGTATATGATGATGCTGA 265  
Db 428 AGAATGATGAGCCATGCTCTAGTGTGAAGACAGAGATGGGTATATGATGATGCTGA 487  
QY 266 TGCAGAGAGAGAGGACTTATCAACCTGGAAGAGTCTCTCATGAGCCAAACAGTGG 325  
Db 488 TGCAGAGAGAGAGGACTTATCAACCTGGAAGAGTCTCTCATGAGCCAAACAGTGG 547  
QY 326 TAAATCTGCAATTTGATTAAGCTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 385  
Db 548 TAAATCTGCAATTTGATTAAGCTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 607  
QY 386 AATGCTGCTTCTATGAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 445  
Db 608 AATGCTGCTTCTATGAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 667  
QY 446 GGTCTGACAGATCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 505  
Db 668 GGTCTGACAGATCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 727  
QY 506 GGTGAG 565  
Db 728 GGTGAG 787  
QY 566 TCAATTAAG 625  
Db 788 TCAATTAAG 847  
QY 626 TGTCTGAG 685  
Db 848 TGTCTGAG 907  
QY 686 GAATCCGATTAATTAAGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 745  
Db 908 GAATCCGATTAATTAAGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 967  
QY 746 AAGCTGCTGCAACAGAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 805  
Db 968 AAGCTGCTGCAACAGAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1027  
QY 806 AGTCAATCTTCTGATGAGAGTGTTCATATATCAAGTATGAGAGAGAGAGAGAGAGAGAG 865  
Db 1028 AGTCAATCTTCTGATGAGAGTGTTCATATATCAAGTATGAGAGAGAGAGAGAGAGAGAG 1087  
QY 866 GCTTCTGCTTAAAG 925  
Db 1088 GCTTCTGCTTAAAG 1147  
QY 926 TGTCTTCAAGATGCAAG 985  
Db 1148 TGTCTTCAAGATGCAAG 1207

QY 986 CAGCTCGGGAGAGTACCTGCTCCGCTATTGAGTCAGGAGCGAAGCTGA 1045  
 DB 1208 AAGCTTTGAGAGCGCTATTATCATCTGCTCTTGAATCCGGAAGCAGAGCAGA 1267  
 QY 1046 AAGCATGACTTTTGAAGCCCTGAATCCCGTTAAGGCTCTCACTACTGAATTTCTGTT 1105  
 DB 1268 GAGCATGCTCTTGAAGCCCTGAATATGAAGTATAGAGCCATTGCTGCTTTCAAGTT 1327  
 QY 1106 ACTT 1109  
 DB 1328 GCAT 1331

RESULT 4

US-09-938-842A-893  
 ; Sequence 893, Application US/09938842A  
 ; Patent No. US20020160378A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harper, Jeff  
 ; APPLICANT: Kreps, Joel  
 ; APPLICANT: Wang, Xun  
 ; APPLICANT: Zhu, Tong  
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
 ; FILE REFERENCE: SCRIP1300-3  
 ; CURRENT APPLICATION NUMBER: US/09/938,842A  
 ; CURRENT FILING DATE: 2001-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/227,866  
 ; PRIOR FILING DATE: 2000-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/264,647  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: US 60/300,111  
 ; PRIOR FILING DATE: 2001-06-22  
 ; NUMBER OF SEQ ID NOS: 5379  
 ; SEQ ID NO 893  
 ; LENGTH: 966  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 ; US-09-938-842A-893

Query Match 43.7%; Score 595.2; DB 9; Length 966;  
 Best Local Similarity 76.5%; Pred. No. 1.2e-157;  
 Matches 748; Conservative 0; Mismatches 218; Indels 12; Gaps 1;

QY 90 ATGCTGTTGAAGTCCGGAATGCCAAGATGTTACGGAATGATTGTAACCCCA 149  
 DB 1 ATGCTGTTGAAGTCCGGAATGCCAAGATGTTACGGAATGATTGTAACCCCA 60  
 QY 150 TTAGTATCTAATAAAGTCCGGAATGCCAAGATGTTACGGAATGATTGTAACCCCA 209  
 DB 61 TTAGTATCTAATAAAGTCCGGAATGCCAAGATGTTACGGAATGATTGTAACCCCA 120  
 QY 210 TTGATGAGCCATGCTCTAGTGTGAAGCAGAGATTGGTATAGTATGATTGCTGATGCA 269  
 DB 121 ATGATGAGCCATGCTCTAGTGTGAAGCAGAGATTGGTATAGTATGATTGCTGATGCA 180  
 QY 270 GAAGAGAGGAGCTTATCAACAGCTGGAAGAGTGTCTCATGAGCCCAAGTGTAT 329  
 DB 181 GAAGAGAGGAGCTTATCAACAGCTGGAAGAGTGTCTCATGAGCCCAAGTGTAT 240  
 QY 330 ACTGCAATGATTAAGCTTCAAGCAGAGCCAGGCTTCAAGCTCATATTAATGATG 389  
 DB 241 ACTGCAATGATTAAGCTTCAAGCAGAGCCAGGCTTCAAGCTCATATTAATGATG 300  
 QY 390 CCGCTTCTATGAGTCTTGAAGAGAGATCATTTCTATTAGCTTTGAGAGCTGATGCT 449  
 DB 301 CCGCTTCTATGAGTCTTGAAGAGAGATCATTTCTATTAGCTTTGAGAGCTGATGCT 360  
 QY 450 CTGACAGATCTGCTAAGGAGATGAAGTGTCTGTTCAAGAGCTGAAGATATTGCT 509  
 DB 361 CTGACAGATCTGCTAAGGAGATGAAGTGTCTGTTCAAGAGCTGAAGATATTGCT 420  
 QY 510 AAGAGCCCAATGCTACATCTTCAACATTTGAAAACCTGCCAATCCCAAGTTGAT 569

DB 421 AAAACACCTATATGTTATATGCTTCAACAGTTTGAGAAATCTCTATATCCAAAGATCCAT 480  
 QY 570 TATGAACCACTGCTTCCAGAGATATGAAGGCTCCGATGGGAAATATGATGCTTTGTT 629  
 DB 481 TATGAACCACTGCTTCCAGAGATATGAAGGCTCCGATGGGAAATATGATGCTTTGTT 540  
 QY 630 TCTGGATAGGAGCTGCTGTTGATCAATTAAGGCTCTGAAATATCTTAAAGAGAGAT 689  
 DB 541 TCTGGATAGGAGCTGCTGTTGATCAATTAAGGCTCTGAAATATCTTAAAGAGAGAT 588  
 QY 690 CCGATATTAAGCTGATTTGTTGGAACCACTGAAAGTCCAGTCTCTAGAGAGAAAG 749  
 DB 589 ACAAACATTAAGCTGATTTGTTGGAACCACTGAAAGTCCAGTCTCTAGAGAGAAAG 648  
 QY 750 CCGTGTCCACAGATTCAGAGGAGTGTGCTGTTTATCCCTGCTGTTGAAAGTC 809  
 DB 649 CCGTGTCCACAGATTCAGAGGAGTGTGCTGTTTATCCCTGCTGTTGAAAGTC 708  
 QY 810 AATCTTCTGATGAAGTGTTCATATATCAAGTATGAAGCAATAGAACTGCAAGCTT 869  
 DB 709 GATCTTATAGTGAAGTGTTCATATATCAAGTATGAAGCAATAGAACTGCAAGCTT 768  
 QY 870 CTGCGCTTAAGAGAGCTTATTTGTTGGAATATCTTCCGAGCTGAGCTGCTGCT 929  
 DB 769 CTGCGCTTAAGAGAGCTTATTTGTTGGAATATCTTCCGAGCTGAGCTGCTGCT 828  
 QY 930 TTTCAGATTCAGAAAGAGAGCAAGAAATGCCGGAAGCTTATTTGCTTTTCCAGC 989  
 DB 829 ATCAACTTGCAGAGAGAGAGCAAGAAATGCCGGAAGCTTATTTGCTTTTCCAGC 888  
 QY 990 TTGCGGAGAGAGTACCTGCTCCGCTGCTTATTTGAGTCAAGTGAAGAGAGAGCTGAAGC 1049  
 DB 889 TTGCGGAGAGAGTACCTGCTCCGCTGCTTATTTGAGTCAAGTGAAGAGAGAGCTGAAGC 948  
 QY 1050 ATGACTTTGAGCCCTGA 1067  
 DB 949 ATGACTTTGAGCCCTGA 966

RESULT 5

US-09-938-842A-893  
 ; Sequence 893, Application US/09938842A  
 ; Publication No. US20040009476A9  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harper, Jeff  
 ; APPLICANT: Kreps, Joel  
 ; APPLICANT: Wang, Xun  
 ; APPLICANT: Zhu, Tong  
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
 ; FILE REFERENCE: SCRIP1300-3  
 ; CURRENT APPLICATION NUMBER: US/09/938,842A  
 ; CURRENT FILING DATE: 2001-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/227,866  
 ; PRIOR FILING DATE: 2000-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/264,647  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: US 60/300,111  
 ; PRIOR FILING DATE: 2001-06-22  
 ; NUMBER OF SEQ ID NOS: 5379  
 ; SEQ ID NO 893  
 ; LENGTH: 966  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 ; US-09-938-842A-893

Query Match 43.7%; Score 595.2; DB 11; Length 966;  
 Best Local Similarity 76.5%; Pred. No. 1.2e-157;  
 Matches 748; Conservative 0; Mismatches 218; Indels 12; Gaps 1;

QY 90 ATGCTGTTGAAGTCCGGAATGCCAAGATGTTACGGAATGATTGTTAAGCCCA 149

Db 1 ATGGCTTCAGTGGCTCCTAAATCGCTAAAGATGTAAGTGAAGTAAAGGAACTCCCA 60  
QY 150 TTATATATCTTAAATTAAGCTGGGATGCTGTGTGCCCCGGGTGCTGTAAGTGGAG 209  
Db 61 TTATATATCTTAAATTAAGCTGGGATGCTGTGTGCCCCGGGTGCTGTAAGTGGAG 120  
QY 210 TTGATGAGCCATGCTCTAGTGTGAAGCAAGATTTGGTATATGATTTGCTGATGCA 269  
Db 121 ATGATGAGCCATGCTCTAGTGTGAAGCAAGATTTGGTATATGATTTGCTGATGCA 180  
QY 270 GAAGAGAGGAGCTTATCAACCTGGAAAAGAGTGTCTCATTTGAGCCCAACAAGTGTAT 329  
Db 181 GAAGCTAAAGAGCTTATTAACCGGAGAGAGTGTGTGATTTGAGCCCAACAAGTGTAT 240  
QY 330 ACTGCGATTGATTTAGCTTCAAGCTGAGAGCCAGGGGTACAGCTCATTAATTACATG 389  
Db 241 ACTGCGATTGATTTAGCTTCAAGCTGAGAGCCAGGGGTACAGCTCATTAATTACATG 300  
QY 390 CCTGCTTTCTAGTGTCTTGAAGAGATCATTTCTATTAGCTTTTGGAGCTGAGTGTGTT 449  
Db 301 CCAGCTTCAATGATATAGAGAGAGATCATTTCTATTAGCTTTTGGAGCTGAGTGTGTT 360  
QY 450 CTGACAGATCTGCTTAAGGAGATGAAGTGTCTGTTCAAGAGCTGAAGAGATATTTGCT 509  
Db 361 CTGACAGATCTGCTTAAGGAGATGAAGTGTCTGTTCAAGAGCTGAAGAGATATTTGCT 420  
QY 510 AAGAGCCCAATGCTTACATCTTCAACAATTGAAAACCTGCAATCCCAAGTGTAT 569  
Db 421 AAGAGCCCAATGCTTACATCTTCAACAATTGAAAACCTGCAATCCCAAGTGTAT 480  
QY 570 TATGAAACCACTGCTCAGAGATATGAAGGCTCCGATGGGAAAATTGATGATTTGTT 629  
Db 481 TATGAAACCACTGCTCAGAGATATGAAGGCTCCGATGGGAAAATTGATGATTTGTT 540  
QY 630 TCTGGATAGGCACTGTGTGTAACAATACAGTGTGGAATAATCTTAAAGAGCAGAT 689  
Db 541 TCTGGATAGGCACTGTGTGTAACAATACAGTGTGGAATAATCTTAAAGAGCAGAT 588  
QY 690 CCGAATATAAGCTGATTTGTGTGAACCACTGTAAGTCCAGTGTCTCAGAGAGAG 749  
Db 589 ACGAATATAAGCTGATTTGTGTGAACCACTGTAAGTCCAGTGTCTCAGAGAGAG 648  
QY 750 CCTGTCCACACAGATTTCAAGGATTTGCTGTGTTTATCCCTGCTGTGTAAGT 809  
Db 649 CCTGTCCACACAGATTTCAAGGATTTGCTGTGTTTATCCCTGCTGTGTAAGT 708  
QY 810 AATCTTCTGATGAGTGTTCATATCAAGTGTGAAGCAATAGAACTGCAAGCTT 869  
Db 709 AATCTTCTGATGAGTGTTCATATCAAGTGTGAAGCAATAGAACTGCAAGCTT 768  
QY 870 CTTGCGCTTAAAGAGAGGCTATTTGTGGGAATATCTTCCGAGCTGAGCTGTGCT 929  
Db 769 CTTGCGCTTAAAGAGAGGCTATTTGTGGGAATATCTTCCGAGCTGAGCTGTGCT 828  
QY 930 TTTGAGATTGCAAAAAGACCAAGAAATGCTGGGAGAGCTTATTTGCTGCTGAGC 989  
Db 829 ATCAAACTTGCAAAAGAGGCTGGAATAATGCTGGGAGAGCTTATTTGCTGCTGAGC 888  
QY 990 TTGCGGAGAGAGTACCTGTCTGCTGCTATTTGAGTCAAGTGAAGAGCGGAGCTGAAGC 1049  
Db 889 TTGCGGAGAGAGTATCTATGAGCTGATCTGTTGATGAGCAAGAGAGAGCGGAGCT 948  
QY 1050 ATGACTTTTGAAGCCCTGA 1067  
Db 949 ATGACTTTTGAAGCCCTGA 966

RESULT 6  
US-10-424-599-55160  
; Sequence 55160, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 55160  
; LENGTH: 1404  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_20820C.1  
US-10-424-599-55160

Query Match 42.5%; Score 578.6; DB 13; Length 1404;  
Best Local Similarity 74.0%; Pred. No. 7.6e-153;  
Matches 759; Conservative 0; Mismatches 264; Indels 2; Gaps 2;

QY 86 TATAATGCTGTTGAAGTCCCGAATTCGCAAGATGTACGGAATGATTTGTAATAAC 145  
Db 147 TATAATGCTGTTGAAGTCCCGAATTCGCAAGATGTACGGAATGATTTGTAATAAC 206  
QY 146 CCCATTAATATCTTAATAAATGCTGCGAT-GGTTGTGTTGCCCGGTTGCTGTAAC 204  
Db 207 ACTGCACTTTTGTGTTTGAAGGAAATTTAAGCAATCTATCTTAACTTGGGAATC 266  
QY 205 TGGAGTTGATGAGCCATGCTCTAGTGTGAAGAGACAGATTTGGTATAGATTTGCTG 264  
Db 267 TAGAATGATGAGCCATGCTCTAGTGTGAAGAGACAGATTTGGTATAGATTTGCTG 326  
QY 265 ATGCAAGAGAGAGGAGCTTATCAACCTGGAAGAGTGTCTCATTTGAGCCCAAGATG 324  
Db 327 ATGCTGAGAGAGAGGAGCTTATCAACCTGGAAGAGTGTCTCATTTGAGCCCAAGATG 386  
QY 325 GTAATCTGCAATGATTTAGCTTCAAGCTGAGAGCCAGGGGTATCAAGCTCATTAATTA 384  
Db 387 GAAACAGGAGAGGAGGTTGGCAATTCATGAGCTGCTTAAGGTTATTAATTAATTA 446  
QY 385 CAATGCTGCTTATAGAGTCTTGAAGAGATCATTTCTATTAGCTTTGAGCTGAGT 444  
Db 447 CAATGCTGCTTATAGAGTCTTGAAGAGATCATTTCTATTAGCTTTGAGCTGAGT 506  
QY 445 TGGTCTGACAGATCTGCTTAAGGAGATGAAGTGTCTGTTCAAGAGCTGAGAGATAT 504  
Db 507 TGGTCTGACAGATCTGCTTAAGGAGATGAAGTGTCTGTTCAAGAGCTGAGAGATAT 566  
QY 505 TGGCTAAGAGCCCAATGCTTCAATATGAAAGGCTGCTGTTCAAGAGCTGAGAGATAT 564  
Db 567 TGGCTAAGAGCCCAATGCTTCAATATGAAAGGCTGCTGTTCAAGAGCTGAGAGATAT 626  
QY 565 TGCATTATGAACCACTGCTTCAAGATATGAAAGGCTGCTGTTCAAGAGCTGAGAGAT 624  
Db 627 TGCATTATGAACCACTGCTTCAAGATATGAAAGGCTGCTGTTCAAGAGCTGAGAGAT 685  
QY 625 TTTGTTCTGAGTGAAGCACTGTTGTAACAATGAGTGTGCTGAAATATCTTAAAGAGC 684  
Db 686 TTTGTTCTGAGTGAAGCACTGTTGTAACAATGAGTGTGCTGAAATATCTTAAAGAGC 745  
QY 685 AGAATCCGAATATAAGCTGATTTGTTGTAACAAGTGTGAAAGTCCAGTGTCTGAGAG 744  
Db 746 AGAATCCGAATATAAGCTGATTTGTTGTAACAAGTGTGAAAGTCCAGTGTCTGAGAG 805  
QY 745 GAAAGCTGTTCAACAAGATTTCAAGGATTTGTTTATCCCTGCTGTTGCTGAG 804  
Db 806 GAAAGCTGTTCAACAAGATTTCAAGGATTTGTTTATCCCTGCTGTTGCTGAG 865  
QY 805 AAGTCAATCTTCTGATGAGTGTTCATTAATATCAAGTATGAGCAATGAAACTGCAA 864  
Db 866 AAGTCAATCTTCTGATGAGTGTTCATTAATATCAAGTATGAGCAATGAAACTGCAA 925



QY 865 AGCTTCTGCGCTTAAGAGGCGCTAATTTGGGAATATCTCCGAGCTGAGCTGCTG 924  
 DB 926 AGCTTCTGCGCTTAAGAGGCGCTAATTTGGGAATATCTCCGAGCTGAGCTG 985  
 QY 925 CTGCTTTCAGATTGCAAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 984  
 DB 986 CTGAGTTAAGATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1045  
 QY 985 CCAGCTTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1044  
 DB 1046 CAAGTTTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1105  
 QY 1045 AAAGCATGACTTTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1104  
 DB 1106 AGAGCTTGTCTTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1165  
 QY 1105 TACTT 1109  
 DB 1166 TGCAAT 1170

RESULT 7  
 US-10-425-114-25435  
 ; Sequence 25435, Application US/10425114  
 ; Publication No. US20040034888A1

GENERAL INFORMATION:  
 APPLICANT: Liu, Jingdong  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Screen, Steven E.  
 APPLICANT: Tabaska, Jack E.  
 APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(5313)B  
 CURRENT APPLICATION NUMBER: US/10/425,114  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 73128  
 SEQ ID NO 25435  
 LENGTH: 1319  
 TYPE: DNA  
 ORGANISM: Zea mays  
 FEATURE:  
 OTHER INFORMATION: Clone ID: LIB3900-005-C7\_FLI  
 US-10-425-114-25435

Query Match 40.9%; Score 557.4; DB 13; Length 1319;  
 Best Local Similarity 72.7%; Pred. No. 7.5e-147;  
 Matches 720; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

QY 75 TAGTACGACATATATATGCTTTGAAAGGCTCCGGAATGCCAAGATGTTACGGAATG 134  
 DB 76 TCGTACGTCGACCAATGAGGAGGCTCGCGCTGATGCCAAGAGCTACCGAGTTG 135  
 QY 135 ATTGTAACCAACCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 194  
 DB 136 ATGCGCAACAGCGCGCTGTGTAATCAACAAGTGAACGATGCGTGGCGCGTC 195  
 QY 195 GCTGCTAACTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 254  
 DB 196 GCGCGCAAGCTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 255  
 QY 255 ATGATTTGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 314  
 DB 256 ATGATTCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 315  
 QY 315 CCAACAAGTGTATATGCTGCTTGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 374  
 DB 316 CCAACTAGCGGCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 375  
 QY 375 CTGATATTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 434  
 DB 376 CTGATATTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 435

QY 435 GAGCTGAGTGTGTTCTGACAGATCTGCTTAAGGAATGAAGGAGTGTGTTCAAGGCT 494  
 DB 436 GGTGCTGACTGTCTTACTGACACCACTCTTGGGAATGAAGAGGCTGTCAAGAGCG 495  
 QY 495 GAAGATATTTGGCTTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 554  
 DB 496 GAAGATATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 555  
 QY 555 AATCCAGGTTCAATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 614  
 DB 556 AACCMAAGATTCAATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 615  
 QY 615 ATTGATGATTTGTTTCTGGATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 674  
 DB 616 ATTGATGAGCTTGTATCTGCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 675  
 QY 675 CTTAAGAGCAGAAATCCGAATATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 734  
 DB 676 CTCAGAGAGCAAAATCTTAATGCTCAAGCTTATGAGGAGGAGGAGGAGGAGGAGG 735  
 QY 735 CTCTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 794  
 DB 736 TTGAATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 795  
 QY 795 GGTGCTTGAAGTCAATCTTCTGATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 854  
 DB 796 GAGTCTTGAAGTGAATCTCTGATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 855  
 QY 855 GAACTGCAAGGCTTCTGCGCTTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 914  
 DB 856 GAGACTGCCAAGGCTCTGCTCTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 915  
 QY 915 GCAGCTGCTGCTTTTTCAGATTGCAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 974  
 DB 916 GCTGCAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 975  
 QY 975 GCGGTTTTTCCAGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1034  
 DB 976 GTGCTGTCCGAGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1035  
 QY 1035 CGGAGGCTGAAAGCATGACTTTGAGGCTT 1065  
 DB 1036 AAGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1066

RESULT 8  
 US-10-425-114-19467  
 ; Sequence 19467, Application US/10425114  
 ; Publication No. US20040034888A1

GENERAL INFORMATION:  
 APPLICANT: Liu, Jingdong  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Screen, Steven E.  
 APPLICANT: Tabaska, Jack E.  
 APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(5313)B  
 CURRENT APPLICATION NUMBER: US/10/425,114  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 73128  
 SEQ ID NO 19467  
 LENGTH: 1308  
 TYPE: DNA  
 ORGANISM: Zea mays  
 FEATURE:  
 OTHER INFORMATION: Clone ID: LIB3150-027-D8\_FLI  
 US-10-425-114-19467

Query Match 40.8%; Score 555.8; DB 13; Length 1308;  
 Best Local Similarity 72.6%; Pred. No. 2.1e-146;

Matches 719; Conservative 0; Mismatches 272; Indels 0; Gaps 0;

```

Oy 75 TAGTGAGCAGATATATATGCTGTGTAAGGTCGGGAATGCCAAGATGTTACGGAATG 134
Db 65 TCGTGAAGTGAACCCCAATGGAAGGCTCGCGTCAATGCGCAAGAGCGTCAACGAGTTG 124
Oy 135 ATTGTTAAACCCCAATGATATATTAATAAATGCGGATGCTGTGTTGCCCGGTT 194
Db 125 ATCGGAAACAGCGCGTGTGTACTCAACAAGGTGACGATGCGTGTGCGCGCGTTC 184
Oy 195 GCTGCTAACTGAGTTGATGAGCGCATCTCTAGTGTGAAGCAGATGGTATAGT 254
Db 185 GCGGCAAGCTCGAGTCCATGAGCGCTGCTCCAGCGTCAAGATAGGATGCTACAGC 244
Oy 255 ATGATGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 314
Db 245 ATGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 304
Oy 315 CCAACAGAGTGTATATCTGCGATTTAGCTTCAATGCGAGCGAGCGGTTACAG 374
Db 305 CCAACTGAGCGCAACAGAGCATTTGAGCTGCTTATGCTGCTGCGCAAGGCTACAA 364
Oy 375 CTCTAATTAACAATGCTGCTCTATGAGTCTTGAGAGAGAGATCTATTAAGCTTT 434
Db 365 CTTACACTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 424
Oy 435 GAGAGTGAAGTGTGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 494
Db 425 GGTGCTGAAGTGTGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 484
Oy 495 GAGAGATATTTGCTTAAGAGCGCGCAATGCTTACATCTTCAACAATTTGAAAAC 554
Db 485 GAGAGATATCAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 544
Oy 555 AATCCCAAGTTCATTAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 614
Db 545 AATCCCAAGTTCATTAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 604
Oy 615 ATGATGCTGATGCTTCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 674
Db 605 ATGCTGCTGCTGATGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 664
Oy 675 CTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 734
Db 665 CTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 724
Oy 735 CTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 794
Db 725 TTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 784
Oy 795 GGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 854
Db 785 GAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 844
Oy 855 GAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 914
Db 845 GAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 904
Oy 915 GCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 974
Db 905 GCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 964
Oy 975 GCGGTTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1034
Db 965 GTCGTGTTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1024
Oy 1035 CGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1065
Db 1025 AAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1055

```

RESULT 9  
US-10-425-114-23298

```

; Sequence 23298, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 23298
; LENGTH: 1312
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3597-017-F4_FLI
US-10-425-114-23298

```

Query Match 40.8%; Score 555.8; DB 13; length 1312;  
Best Local Similarity 72.6%; Pred. No. 2.1e-146;  
Matches 719; Conservative 0; Mismatches 272; Indels 0; Gaps 0;

```

Oy 75 TAGTGAGCAGATATATATGCTGTGTAAGGTCGGGAATGCCAAGATGTTACGGAATG 134
Db 69 TCGTGAAGTGAACCCCAATGGAAGGCTCGCGTCAATGCGCAAGAGCGTCAACGAGTTG 128
Oy 135 ATTGTTAAACCCCAATGATATATTAATAAATGCGGATGCTGTGTTGCCCGGTT 194
Db 129 ATCGGAAACAGCGCGTGTGTACTCAACAAGGTGACGATGCGTGTGCGCGCGTTC 188
Oy 195 GCTGCTAACTGAGTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 254
Db 189 GCGGCAAGCTCGAGTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 248
Oy 255 ATGATGCTGATGCTTCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 314
Db 249 ATGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 308
Oy 315 CCAACAGTGTATATCTGCGATTTAGCTTCAATGCGAGCGAGCGGTTACAG 374
Db 309 CCAACTGAGCGCAACAGAGATTTGAGTGGCTTATGCTGCTGCGCAAGGCTACAA 368
Oy 375 CTCTAATTAACAATGCTGCTCTATGAGTCTTGAGAGAGAGATCTATTAAGCTTT 434
Db 369 CTTACACTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 428
Oy 435 GAGAGTGAAGTGTGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 494
Db 429 GGTGCTGAAGTGTGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 488
Oy 495 GAGAGATATTTGCTTAAGAGCGCGCAATGCTTACATCTTCAACAATTTGAAAAC 554
Db 489 GAGAGATATCAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 548
Oy 555 AATCCCAAGTTCATTAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 614
Db 549 AATCCCAAGTTCATTAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 608
Oy 615 ATGATGCTGATGCTTCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 674
Db 609 ATGATGCTGCTGATGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 668
Oy 675 CTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 734
Db 669 CTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 728
Oy 735 CTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 794

```

Db	729	TTGAATGCTGAAAACCTTGACCCACACAAAGATTCAAGGAATTGAGCTGCTTTATCCCT	788
QY	795	GGTGCTTGGAGTCATCTCTTGATGAAGTTGTTCAAAATATCAAGTGATGAAGCAATA	854
Db	789	GGAGTCTTGATGTGATCTCTTGATGAAACTCTACAGTTTCAAGTGATGAAGCTATC	848
QY	855	GAAACTGCAAGCTTCTGCGCTTAAAGAAGGCCATATTGCGAATATCTTCCGAGCT	914
Db	849	GAGACTGCCAAGGCTCTTGCTCTGAAGAAGGTTGTTGGTAATCTCTTGATGCA	908
QY	915	GCACTGCTGCTGCTTTTCAGATTGCCAAAAGACCAGAAAATGCCGGGAAGCTTATTGT	974
Db	909	GCTGCAGCTGCGGCAGTTAGGCTTGTCTTAGAGGCCAGAAAACCGCGAAGCTATTGTT	968
QY	975	GCCGTTTTTCCAGCTTCGGGAGAGGTAACCTGTCCCTCCGTGCTATTGAGTCAAGTGA	1034
Db	969	GTCGTGTTCCCGAGCTTCGGCGAGCGCTACCTCTCATCGTGTGTTCCAGTCCATCAAG	1028
QY	1035	CGCGAAGCTGAAAGCATGACTTTGAGCCCT	1065
Db	1029	AAGGAAGCGGAGAGCATGTGTGCGAGCCCT	1059

```

RESULT 10
US-10-425-114-1445
; Sequence 1445, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 1445
; LENGTH: 1403
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700153393_FLI
US-10-425-114-1445

```

Query Match	40.8%;	Score 555.8;	DB 13;	Length 1403;
Best Local Similarity	72.6%;	Pred. No. 2.2e-146;		
Matches 719; Conservative	0;	Mismatches 272;	Indels 0;	Gaps 0;

OY		75	TAGTGACGAGATATAATGGCTGTGTGAAGAAGSTCCGGAAATTGCCAAGATGTTTACCGAATTG	134
Db		160	TCGTGACGTGACCCCAATGGGAGAGAGCCTCGCCGTCGATCGCCAAGACGTCAACCGAGTTG	219
OY		135	ATTGGTAHAACCCCATTTAGTATATCTAAATTAACCTGCGGATGGCTGTGTGTCGCCCGGTT	194
Db		220	ATCGGGAACAAGCCCCGTGTGTACTCAACAAGGTGACCgATGGTGCCTCGGCCGCTC	279
OY		195	GCTGCTAAACTGGAGTTTGATGAGACCATGCTCTAGTGTGAAGCACAGATTGGGTATAGT	254
Db		280	GCCCCCAAGCTCGAGTCCATGAGCCCTGCTCCAGCGCTCAAGCATAGGATTGGCTACAGC	339
OY		255	ATGATTGCTGATGCAGAAGAGAAAGGGACTTATCACACCTGSAAGAGAGTGTCCTCATTTAG	314
Db		340	ATGATCAAGGACGACAGAGAGAGAAAGGCTGATTACTCCAGCGCTGAGTGTTCTGATTGAA	399
OY		315	CCAACAAGTGTATAACTGTGSCATTGGATTAGCCTTCATGGGACAGCCAGGGGTTACAAG	374
Db		400	CCAAGTAGCGGCACAACAAGGCAITGGACTGGCCTTTATGGCTGTGCTGCCAAGGGCTACAAA	459
OY		375	CTCATTAATTACAATGCTGCTCTCTATGAGTCTTGAGAGAAAGATCATTTCTATTAGCTTTT	434

Db	460	CTTACACTCAAAATGCTGCTCCATGAGCATGAGAGAGATCATATTGAAGCTTTT	519
QY	435	GGAGCTGAGTTGGTTCTGACAGATCCTGTAAAGGGAATGAAGAGTGTGTTCAGAAGCT	494
Db	520	GGTGCTGAACTTGTCTTACTGACCCACTCCTCGGAATGAAGAAGCTGTCAAGAAAGCG	579
QY	495	GAAAGATATTGGCTAAGAAGCCCAATGCCCTACTACTTCAACAATTTGAAAAACCTGCC	554
Db	580	GAAAGATACAAAGCAAGAACAACCAACTCGTACTCTTCAACAATTTGAAAAATCCAGCT	639
QY	555	AATCCCAAGTTTCATTTATGAAACCACTGCTCCAGATATGGAAGAAGCTCCGATGGAAA	614
Db	640	AACCCAAGATTCACTATGAGACTACTGCGCTGAAATCTGGAAGAAGCTACAGCAGGAATA	699
QY	615	ATTGATGCAATTGTTCTTCTGGATAGGCACTGCTGTGTAATTAACAAGTGTGAAAATAT	674
Db	700	ATTGCTGGCTTGATATCTGATCGGACAGAGAGTACATCACCGTACTGCGCGATAC	759
QY	675	CTTAAGAGCAGATCCGAATATAAAGCTGATGTTGTTGTAACCAAGTTGAAGTCCAGTG	734
Db	760	CTCAGAGAGCAAAATCTTAATGTCAAGCTCTATGTTGAGCCAGTGAAGTGTGTT	819
QY	735	CTCTCAGAGGAAGCTTGTCCACACAAGATTCAAGGATTTGCTGTGTTTATCCCT	794
Db	820	TTGAATGTTGAAAACCTGGAACCAACAAGATTCAAGGAATTTGAGCTGTATTATCCCT	879
QY	795	GGTGTCTTGAAGTCAATCTTCTGATGAGAGTTGTTCAATATCAAGTATGAAGCAATA	854
Db	880	GGAGTCTTGAGATGTGATCTCCTTGATGAGACTCTACAGGTTCAAGTATGAAGCTATT	939
QY	855	GAAACTGCAAAAGCTTCTTCGCTTAAGAAGGCTTATTGTGCGGAATATCTCCGAGCT	914
Db	940	GAGACTGCCAAGGCTCTTGCTCTGAAGAAGGGTGTGTTGTAATCTCTTGTGTGCA	999
QY	915	GCAGCTGCTGCTCTTTTTCAGATTGCAAAAAAGCAGAAATGCCGGGAAGCTTATGTT	974
Db	1000	GCTGCAAGCTGCGCAGTTAAGCTTGTCTTAAGAGGCCAGAAAAAGCGGAAAAGCTATTGTT	1059
QY	975	GCCGTTTTTCCAGCTTCGGGGAGAGGTACTGTCTCCGTCTATTGAGTCAGTGGA	1034
Db	1060	GTCGTGTTCCCGAGCTTCGGCGAGCGCTACTCTTCATCGGTGCTGTTCCAGTCCATCAG	1119
QY	1035	CGCGAAGCTGAAGCATGACTTTTGAGCCCT	1065
Db	1120	AAGGAAGCGGAAGCATGTGTGAGCCCT	1150

```

RESULT 11
US-10-425-114-23608
; Sequence 23608, Application US/10425114
; Publication NO. US20040034888A1
GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack B
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 23608
; LENGTH: 1301
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3597-067-G10_FLI
US-10-425-114-23608

```



Query Match 40.7%; Score 554.2; DB 13; Length 1301;  
 Best Local Similarity 72.5%; Pred. No. 6e-146;  
 Matches 718; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

QY 75 TAGTGAGCAGATATATATGCTGTTGAAAGTCCGGAATTGCCAAGATGTACGGAATTG 134  
 DB 58 TCGTGACGTGACCCCAATGGGAGAGGCTCGCCGTGATCGCCCAAGGACGTACCGAGTTG 117  
 QY 135 ATTGGTAAACCCCATTTAGTATATCTAAATTAACCTTGGGATGTTGTGTCGGGGTT 194  
 DB 118 ATCGGAACAGCCCGCTGTTGTTACCTCAACAGGTGACCGATGGTGGTGGCCACGTC 177  
 QY 195 GCTGTAACCTGAGTGTGATGAGCCATGCTCTAGTGTGAAGACAGATTGGTATAGT 254  
 DB 178 GCGGCAAGCTCGAGTCCATGAGCCCTGCTCCAGCGTCAAGATAGATTGGCTACAGC 237  
 QY 255 ATGATGCTGATGAGAGAGAGAGAGGACTTATCAACCTGGAAGAGTGTCTCATTTAG 314  
 DB 238 ATGATCAAGAGCAGAGAGAGAGAGGCTGATTAATCTCAGGCGTGAATGTTCTGATTGA 297  
 QY 315 CCAACAGTGTATATCTGCGATTTAGCTTCACTGAGCAGAGCCAGGGTTACAAG 374  
 DB 298 CCACTAGCGGCAACAGAGGATTTGAGCTGCTTATGCTGCTGCCAAGGCTACAA 357  
 QY 375 CTGATATTAATATGCTGCTGCTTCTATGATCTTGAAGAGAGATCATTTCTATTAGCTTT 434  
 DB 358 CTTACACTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 417  
 QY 435 GAGCTGAGTGTGCTTCTGACAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 494  
 DB 418 GGTGCTGAACTTGTCTTACTGACCCACTCTTGGGATGAAGAGCTGTCAAGAAAGCG 477  
 QY 495 GAAGAGATATTTGCTTGAAG 554  
 DB 478 GAAGAGATATCAAG 537  
 QY 555 AATCCCAAGTCTTATTAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 614  
 DB 538 AATCCCAAGTCTTATTAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 597  
 QY 615 ATTGATGATTTGTTCTGAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 674  
 DB 598 ATTGATGCTTGTATCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 657  
 QY 675 CTTAAGAGAGAGATCCGATATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 734  
 DB 658 CTTAAGAGAGAGATCCGATATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 717  
 QY 735 CTCTCAG 794  
 DB 718 TTGAATGTTGAAAACTGAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 777  
 QY 795 GGTGCTTGAAGTCAATCTTCTGATGAAGTGTTCATATATCAAGTGAAGCAATA 854  
 DB 778 GGAGTCTGATGATGATCTCTGATGAAGTGTTCATATATCAAGTGAAGCAATA 837  
 QY 855 GAACTGCAAGCTTCTGCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 914  
 DB 838 GAGACTGCAAGCTTCTGCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 897  
 QY 915 GAGCTGCTGCTCTTCTGATTTGCAATTTGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 974  
 DB 898 GCTGAGCTGCGGAGTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 957  
 QY 975 GCGGTTTCCAGCTTGGGAG 1034  
 DB 958 GTGTGTTCCAGCTTGGGAG 1017  
 QY 1035 GCGGAGCTGAAGCATGACTTTGAGCCCT 1065  
 DB 1018 AAGGAGCGGAG 1048

RESULT 12  
 US-10-424-599-61234  
 ; Sequence 61234, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa Thomas J  
 ; APPLICANT: Kovalic David K  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 61234  
 ; LENGTH: 1603  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_26307C.1  
 ; US-10-424-599-61234

Query Match 39.9%; Score 543.4; DB 13; Length 1603;  
 Best Local Similarity 73.1%; Pred. No. 7.9e-143;  
 Matches 697; Conservative 0; Mismatches 256; Indels 0; Gaps 0;

QY 108 GGAATTCGCCAAGATGTTACGGAATGATTTGTTAAACCCCATTTAGTATATCTAAATAA 167  
 DB 233 GCAATCAAGAGAGATGCCATGAAATTTGTTAGGAAATCCCAATGTTATTTCTAAACAAC 292  
 QY 168 CTGCGGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 227  
 DB 293 ATTGTTGCTGGGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 352  
 QY 228 ACTGTGAAGAGAGAGATTTAGTATGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 287  
 DB 353 AGTGTCAAGATGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 412  
 QY 288 AACCTGGAAGAGAGTGTCTCATTTAGAGCCCAAGTGTATATCTGCAATTGATTTAGCC 347  
 DB 413 ACTCTGGAAGAGAGTGTCTCATTTAGAGCCCAAGTGTATATCTGCAATTGATTTAGCC 472  
 QY 348 TTCAATGAG 407  
 DB 473 TTCAATGAG 532  
 QY 408 GAGAGAGAGATCATTTCTTATTTAGCTTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 467  
 DB 533 GAGAGAGAGATCATTTCTGAG 592  
 QY 468 GGAATGAAGAGTGTCTGTTGAG 527  
 DB 593 GGAATGATGAGATTTATGAG 652  
 QY 528 ATACTCAACATTTGAAACCTTCCCAATCCCAAGTTCTATTTGAAACCACTGTTCCA 587  
 DB 653 ATGCTTAACCAATTTGAAATCTGCAATCCCAATCCCAATTTCTATTTGAAACCACTGTTCCA 712  
 QY 588 GAGATATGAAGAGCTCCGATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 647  
 DB 713 GAGATTTGAGAGAGCTCCGAG 772  
 QY 648 GGTACATATACAGTGTCTGGAAG 707  
 DB 773 GGTTCATATGAG 832  
 QY 708 GGTGTGAAG 767  
 DB 833 GGTGTGAAG 892  
 QY 768 CAAGGATTTGTTGTTTATTTCCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 827

Db	893	CAAGGAATTGCTGCTGCTATCATTC	CCAAAGTTCCTGGATGTAATCTC	CTAGATGAAGTT	952
Qy	828	GTTCAATATCAAGTGAATGAAGCAAT	GAAACTGCAAGCTTCTTGCGCTT	AAAGAGGC	887
Db	953	ATTCAAGTTTCAAGTGAAGAGCTAT	AGAACTGCTAACTGCTTGCCCTG	AAAGAGGT	1012
Qy	888	CTATTGTGGAATATCTTCCGGAAGCT	GCAGCTGCTGCTGCTTTCAGATTG	CAAAAGA	947
Db	1013	TTGCTGATGGGAATTTCA	TGAGAGCTGCTGCAGCAGCAAT	AAAGTGGGAGAGA	1072
Qy	948	CCAGAAATGCCCCGGAAGCTTAT	TGTCGGTTTTCCAGCTTCCGGGAGAG	GTACTG	1007
Db	1073	CCAGAAATGCAAGAAAGCTCAT	TGTTGTGATTTTCCCAAGTTTGGA	GAACGTACTTA	1132
Qy	1008	TCCTCCGTGCTATTGAGTCA	GTGAGACCGGAAGCTGAAAGCAT	GACTTTGA	1060
Db	1133	TCTTCACCACTGTTGATTC	CAATTAGGAAAGAGAGCTGAACA	AAATGACATTGA	1185

```

RESULT 13
US-10-425-114-22525
; Sequence 22525, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingsong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 22525
; LENGTH: 1279
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3587-229-H7_F11
US-10-425-114-22525

```

Query Match	38.5%	Score 524.4	DB 13	Length 1279
Best Local Similarity	74.2%	Pred. No. 1.6e-137		
Matches 663	Conservative 0	Mismatches 231	Indels 0	Gaps 0
QY	172	CGGATGTTGTGTTGCCCGGCTGCTCTAACTGAGTTGATGAGCCATGCTTACTG	231	
Db	133	CCGATGGTGCGTGGCCGGCTGCCGCCAAGCTCGAGTCCATGGAGCCCTGCTCAGCG	192	
QY	232	TGAAGACAGATTGGGTATAGTATGATTGCTGATGACAGAGAAGGACTTATCACAC	291	
Db	193	TCAAGATAGATTGGCTACGACATGATCACGGACGACAGAGAGAAGGGCTGATTACTC	252	
QY	292	CTGGAAGAGTGTCTCATTTGACCCCAACAAGTGTATATCTGCATTGGATTAGCCTTCA	351	
Db	253	CAGCGGTAGTGTCTGATTGAATCCAACTAGCGGCAACACAGGCATTGGACTGGCCTTTA	312	
QY	352	TGGCAGCAGCCAGGGGTTACAAGCTCATTAATTACAATGCTCTTCTATAGTCTTGAGA	411	
Db	313	TGGCTGCTGCCAAGGGCTACAACTTACACTCACAAATGCTCTCTCCATGAGCATGAGA	372	
QY	412	GAAGAATCATTTCTATTAGCTTTTGGAGCTGAGTTGTTCTGACAGATCCTGCTAAGGAA	471	
Db	373	GGAGATCATATTGAAGGCTTTTGGTGTGAAGTGTCTCTTAAGTGAACCACTCCTGGAA	432	
QY	472	TGAAGGTGCTGTTCAGAAAGGCTGAAGAGATATTGGCTAAGACGCCCAATGCTTACATAC	531	
Db	433	TGAAGGAGCTGTCAAGAAAGCGAAGAGATACAAAGCAAGAACCCCACTCGTACATCC	492	
QY	532	TTCAACATTTGAAACCCCTGCCAATCCCAAGTTCATTATTAAGAACCACTGCTCCAGAGA	591	

Db	493	TTCAACAATTGAAAATCCAGCTAACCCAAAGTTCACTATGAGACTACTGGGCTGAAA	552
QY	592	TATGGAAGGCTCCGATGGGAAAATTGATGCATTGTTTCTGGATAGCACTGGTGSTA	651
Db	553	TCTGAAAAGCTACGACGAGAAAATTGCTGGGCTTGATCTGTATCGGACAGAGSTA	612
QY	652	CAATACAGGTGCTGGAAAATATCTTAAGAGCAATCCGATATATAAGCTGATTGGTG	711
Db	613	CCATCACCGGTACTGGCGGATACCTCAGAGAGCAAAATCCTAATGTCAAGCTCTATGGTG	672
QY	712	TGGAACCAAGTTGAAAGTCCAGTGTCTCAGAGAGCAAGCCTGTCCACACAAGATTCAAG	771
Db	673	TGAGCCCAAGTGGAGAGTGTCTGTTTGAATGGTGAAGAAACCTGACACACAAGATTCAAG	732
QY	772	GGATTGTGCTGGTTTTATCCCTGTGTCTTGAAGTCAATCTTCTTGATGAAGTTGTTT	831
Db	733	GAATTGAGCTGGTTTTATCCCTGAGTCTTGAATGTGATCTCCTTGATGAAGTCTAC	792
QY	832	AAATATCAAGTGAAGCAATAGAACTGCAGAACTTCTTGCGCTTAAGAGAGGCTAT	891
Db	793	AGGTTCAAGTGAAGCACTATTGAGACTGCCAAGGCTCTTGCTCTGAAAGAGGGTGT	852
QY	892	TTGTGGGAATATCTTCCGAGCTGCAGCTGTGTCTTTCAGATTGCAGAAAAGACCAG	951
Db	853	TGTTGGAAATCTTCTGTGTGCAGCTGCAGCTGCCGCACTTAGGCTTGCTAAGAGGCCAG	912
QY	952	AAATGCCGGGAAGCTTATGTGTGCGGTTTTTCCAGCTTCGGGGAGAGGTAAGTCTT	1011
Db	913	AAAACGCCGGAAGCTATTGT	972
QY	1012	CCGTGCTATTGAGTCACTGAGACGCGAAGCTGAAGCATGACTTTGAGGCCCT	1065
Db	973	CGTGTCTTCCAGTTCATCAGAGGAGAGCGGAAAGCATGTGTGTGTGTGTGTGTGTGTGT	1026

```

RESULT 14
US-10-425-114-34372
; Sequence 34372, Application US/10425114
; Publication NO. US20040034888A1
;
GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 34372
; LENGTH: 1490
;
TYPE: DNA
;
ORGANISM: Zea mays
;
FEATURE:
;
OTHER INFORMATION: Clone ID: UC-ZMFLMO17258D02_FLI
US-10-425-114-34372

```

	Query Match	38.3%;	Score 521.8;	DB 13;	Length 1490;
	Best Local Similarity	71.6%;	Pred. No. 9.9e-137;		
	Matches 685;	Conservative 0;	Mismatches 272;	Indels 0;	Gaps 0;
QY	111 ATTGCCAAGATGTTACGGATTGATTGGTAAACCCATTAGTATATCTAAATAACTT	170			
Db	240 ATGCCAAGACGCTGACCCAGCTCATCCGCAACACACCAATGCTGTATCTCAACACGTC	299			
QY	171 GCGGATGCTGTGTTGCCCCGGGTGCTGCTAAACTGGAAGTTGATGGAGCCATGCTCTAGT	230			
Db	300 GTCAAGGATCTGTGCCCAATGTGCTGCTTAAGCTTCGAGATTATGGAGCCCTGCTTAGC	359			

```
QY 231 GTGAAGACAGATGGGTATAGTATGATGCTGATGAGAGAGAGAGGAGCTTATCACA 290
    |||||||
Db 360 GTGAAGACAGATAGGATAGGATAGATATGATGCTGAGACAGAGAGGCTTATCACA 419
QY 291 CCTGGAAGAGTGTCTCATGAGCCACAGAGTGTATATCTGGCATTTGATGAGCTTC 350
    |||||||
Db 420 CCTGGAAGAGTGTCTCATGAGCCACAGAGTGTATATCTGGCATTTGATGAGCTTC 479
QY 351 ATGGACAGCAGCGGTTACAGCTCATATTAATCAATGCTGCTTCTATGAGCTTGAG 410
    |||||||
Db 480 ATGGCTCTTCTAAGGATATAGCTGATATCAATGCTTCTATGAGCTTGAG 539
QY 411 AGAAGATCATTTCTATTTAGCTTTTGAAGCTGAGTTGTTTGAAGATCTGCTAAGGA 470
    |||||||
Db 540 AGAAGATCTCTCTCAGAGCTTTTGTGCTGAACTTGTCTTACTGATGCTGCAAAAGG 599
QY 471 ATGAAAGTGTCTGTCAGAGGCTGAGAGATATTTGGCTAAGACGCCAATGCTTACATA 530
    |||||||
Db 600 ATGAAAGGCGCTTAGATAGGCTACAGAGATTTTAAACAGACACCAATTTCTTACATG 659
QY 531 CTTCAACATTTGAAAACCTGCGCAATCCCAAGTTCAATATGAAAACCACTGCTCCAGAG 590
    |||||||
Db 660 CTTCAACAGTTGATTAACCTGCGCAACCTTAAGTACATTAATGAGACTACTGCTCCAGAG 719
QY 591 ATATGAAAGGCTCCGATGGAGAAATGATGATTTGTTTCTGGATAGGCACTGCTGCT 650
    |||||||
Db 720 ATCTGGAGAGATTCAAGGAGAGAGTGTATATTCATTTGTTGATTTGAAAGAGGAGG 779
QY 651 ACAATTAACAGGTGCTGAAAATATCTTAAAGACAGATCCGATATTAAGCTGATGCT 710
    |||||||
Db 780 ACAATATCTGCTGCGCGCTTTCTCAAGAGAGAAAATCTGCAATTAAGGTTATGCT 839
QY 711 GTGAAACAGTGAAGTCAAGTCTCTCAGAGAGAAAGCTGCTCCACACAGATTTCAA 770
    |||||||
Db 840 ATTGAGCTTCTGAAGTAACTACTCTCCGTTGAAAACCTGCTCCACATAGATCCAG 899
QY 771 GGGATTGTGCTGTTTATATCCCTGCTGCTTGAAGTCAATCTTCTGATGAGTGT 830
    |||||||
Db 900 GGAATCGGCGAGATTTGTTCCAGAACTTGAAGTCAATCTTCTGATGAGTGT 959
QY 831 CAAATATCAAGTATGAGCAATAGAACTGCAAGCTTCTGCGCTTAAAGAGGCTTA 890
    |||||||
Db 960 GAGATATCAAGTATGAGCTGTTGAGACAGCAAAAGATGCTGTTCAAGAAAGATTA 1019
QY 891 TTTGTGGAATATCTCCGAGCTGAGCTGCTGCTTTTCAAGTTGCAAAAAGACCA 950
    |||||||
Db 1020 CTGCTCGAATCTCTCTGCGGAGCAGACCGCGCGCCATTAAGTTGCCAAAAGACCA 1079
QY 951 GAAAATGCGGAGAGCTTATGTTGCGGTTTCCAGCTTCCGAGAGTACCTGCTC 1010
    |||||||
Db 1080 GAGAATGCTGGAAGCTGATAGTGTGCTGTTCCAGCTTCCGAGAGTACCTTCA 1139
QY 1011 TCCGTCTATTTGAGTCAAGTGAAGAGCGGAGCTGAAGCATGACTTTGAGCCCTGA 1067
    |||||||
Db 1140 TCTGTCTCTATCACTCAATAGAGAAAGATGCGAGAAATGCAAGCTGAGCCATGA 1196
```

## RESULT 15

US-10-425-114-34527

; Sequence 34527, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jindong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven B

; APPLICANT: Tabaska, Jack B

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(5313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

```
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 34527
; LENGTH: 1491
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17281G01_FLI
US-10-425-114-34527
```

```
Query Match          38.3%; Score 521.8; DB 13; Length 1491;
Best Local Similarity 71.6%; Pred. No. 9.9e-137;
Matches 685; Conservative 0; Mismatches 272; Indels 0; Gaps 0;
```

```
QY 111 ATGCCAAAGATGTTACGGAATTGATGTTAAACCCATTAGTATATCTAATAAATCT 170
    |||||||
Db 241 ATGCCAAAGATGTTACGGAATTGATGTTAAACCCATTAGTATATCTAATAAATCT 300
QY 171 GCGATGTTGTTGCTCCGAGTGTGCTTAACTGAGTTGATGAGGAGCTGCTAGT 230
    |||||||
Db 301 GTCAAGGATCTGTGCGCAATGTGCTGCTAAGCTGAGATTATGAGGCTGCTAGC 360
QY 231 GTGAAGACAGATTTGGTATAGTATGATGCTGATGAGAGAGAGAGGACTTATCACA 290
    |||||||
Db 361 GTGAAGACAGATTTGGTATAGTATGATGCTGATGAGAGAGAGGCTTATCACA 420
QY 291 CCTGGAAGAGTGTCTCATTTAGGAGCAAGTGTATATCTGGAATGAGCTTTC 350
    |||||||
Db 421 CCTGGAAGAGTGTCTCATTTAGGAGCAAGTGTATATCTGGAATGAGCTTTC 480
QY 351 ATGGACAGCAGCGGTTACAGCTCATATTAATCAATGCTGCTTCTATGAGCTTGAG 410
    |||||||
Db 481 ATGGCTCTTCTAAGGATATAGCTGATATCAATGCTTCTATGAGCTTGAG 540
QY 411 AGAAGATCATTTCTATTTAGCTTTTGAAGCTGAGTTGTTTGAAGATCTGCTAAGGA 470
    |||||||
Db 541 AGAAGATCTCTCTCAGAGCTTTTGTGCTGAACTTGTCTTACTGATGCTGCAAAAGG 600
QY 471 ATGAAAGTGTCTGTTCAAGAGCTGAGAGATATTTGGCTAAGAGGCTGCTAAGGA 530
    |||||||
Db 601 ATGAAAGGCGCTTAGATAGGCTACAGAGATTTTAAACAGACACCAATTTCTTACATG 660
QY 531 CTTCAACATTTGAAAACCTGCGCAATCCCAAGTTCAATATGAAAACCACTGCTCCAGAG 590
    |||||||
Db 661 CTTCAACAGTTGATTAACCTGCGCAACCTTAAGTACATTAATGAGCTACTGCTCCAGAG 720
QY 591 ATATGAAAGGCTCCGATGGAGAAATGATGATTTGTTTCTGGATAGGCACTGCTGCT 650
    |||||||
Db 721 ATCTGGAGAGATTCAAGGAGGAGTGTATATTCATTTGTTGATTTGGAATGGAACAGGGGG 780
QY 651 ACAATTAACAGGTGCTGAAAATATCTTAAAGACAGAAATCCGATATTAAGCTGATGCT 710
    |||||||
Db 781 ACAATATCTGCTGCGCGCTTTCTCAAGAGAGAAAATCTTGAATTAAGGTTATGCT 840
QY 711 GTGAAACAGTGAAGTCAAGTCTCTCAGAGAGAAAGCTGCTCCACACAGATTTCAA 770
    |||||||
Db 841 ATGAGCTTCTGAAAAGTAACTACTCTCCGTTGAAAACCTGCTCCACATTAAGATCCAG 900
QY 771 GGGATTGTGCTGTTTATATCCCTGCTGCTTGAAGTCAATCTTCTGATGAGTGT 830
    |||||||
Db 901 GGAATCGGCGAGATTTGTTCCAGAACTTGAAGTCAATCTTCTGATGAGTGT 960
QY 831 CAAATATCAAGTATGAGCAATAGAACTGCAAGCTTCTGCGCTTAAAGAGGCTTA 890
    |||||||
Db 961 GAGATATCAAGTATGAGCTGTTGAGACAGCAAAACAGTTGCTGTTCAAGAGGATTA 1020
QY 891 TTTGTGGAATATCTTCCGAGCTGAGCTGCTGCTTTTCAAGTTGCAAAAAGACCA 950
    |||||||
Db 1021 CTGCTCGAATCTCTCTGCGGAGCAGACCGCGCGCCATTAAGTTGCCAAAAGACCA 1080
QY 951 GAAAATGCGGAGAGCTTATTTGCTGCTTTTCCAGCTTCCGAGAGAGGATCTGCTC 1010
    |||||||
Db 1081 GAGAATGCTGGAAGCTGATAGTGTGCTGTTTCCAGCTTCCGAGAGAGGATCTTCA 1140
```



Mon May 24 08:18:49 2004

us-09-931-457a-30.mpb

**Page 12**

```
Oy      1011 TCCCTGCTATTGTAGTCAGTGCAGCGCAAGCTGAAGCATGACTTTTGACCCCTGA    1067
          ||| ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db      1141 TCTGTCTCTATCAGTCCCATAGAAGAAATGCCAGAACATGCAGCCTGAGCCATGA    1197
```

Search completed: May 21, 2004, 22:15:45  
Job time : 707 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2004, 13:00:59 ; Search time 60 Seconds  
(without alignments)  
1530.466 Million cell updates/sec

Title: US-09-931-457A-31  
Perfect score: 1623  
Sequence: 1 MAVERSGIAKQVTLGKTP.....LSVLFESVRREAESMTFEP 325

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: geneseqp1980s:\*

2: geneseqp1990s:\*

3: geneseqp2000s:\*

4: geneseqp2001s:\*

5: geneseqp2002s:\*

6: geneseqp2003as:\*

7: geneseqp2003bs:\*

8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1623	100.0	325	2	AAW81018	Aaw81018 Fragment
2	1431	88.2	325	2	AAR49830	Aar49830 Cysteine
3	1384	85.3	322	3	AAQ29607	Aag29607 Arabidops
4	1384	85.3	322	5	ABR94000	Abb94000 Herbicida
5	1384	85.3	322	5	AAE14774	Aae14774 Arabidops
6	1379	85.0	322	5	ABR94010	Abb94010 Herbicida
7	1379	85.0	322	5	AAE14775	Aae14775 Arabidops
8	1368	84.3	325	7	ABM73727	Abm73727 DNA clone
9	1368	84.3	325	7	ABM73727	Abm73727 DNA clone
10	1296	79.9	305	5	ABR93998	Abm73998 Herbicida
11	1271	78.3	383	2	AAE63756	Aae63756 Spinach C
12	1261	77.7	392	3	AAQ09478	Aag09478 Arabidops
13	1255	77.3	442	5	ABR94007	Abb94007 Herbicida
14	1252	77.1	315	5	ABR94013	Abb94013 Herbicida
15	1252	77.1	315	5	ABR94001	Abb94001 Herbicida
16	1241	76.5	285	3	AAQ29608	Aag29608 Arabidops
17	1239	76.3	284	3	AAQ29609	Aag29609 Arabidops
18	1229	75.7	324	5	ABR94016	Abb94016 Herbicida
19	1229	75.7	324	5	ABR94002	Abb94002 Herbicida
20	1229	75.7	324	5	AAE14776	Aae14776 Arabidops
21	1215	74.9	305	3	AAQ09479	Aag09479 Arabidops
22	1196	73.7	323	3	AAQ25438	Aag25438 Arabidops
23	1192.5	73.5	424	5	ABR94012	Abb94012 Herbicida
24	1192.5	73.5	424	5	ABR93997	Abb93997 Herbicida
25	1192	73.4	323	3	AAQ26543	Aag26543 Arabidops

26	1192	73.4	323	3	AAQ38510	Aag38510 Arabidops
27	1192	73.4	323	5	ABR94014	Abb94014 Herbicida
28	1192	73.4	332	3	AAQ38509	Aag38509 Arabidops
29	1191	73.4	323	3	AAQ44511	Aag44511 Arabidops
30	1191	73.4	333	3	AAQ44510	Aag44510 Arabidops
31	1184	73.0	324	3	AAQ06648	Aag06648 Arabidops
32	1179	72.6	324	3	AAQ35892	Aag35892 Arabidops
33	1179	72.6	324	5	ABR94009	Abb94009 Herbicida
34	1179	72.6	324	5	AAE14779	Aae14779 Thalecres
35	1179	72.6	324	7	ADC64218	Adc64218 Thalecres
36	1179	72.6	399	5	ABR94003	Abb94003 Herbicida
37	1179	72.6	399	5	ABR94005	Abb94005 Herbicida
38	1172	72.2	392	5	ABR94006	Abb94006 Herbicida
39	1149	70.8	323	3	AAQ36615	Aag36615 Arabidops
40	1145	70.5	304	3	AAQ09480	Aag09480 Arabidops
41	1142	70.4	284	3	AAQ26544	Aag26544 Arabidops
42	1142	70.4	304	3	AAQ26544	Aag26544 Arabidops
43	1142	70.4	304	3	AAQ38511	Aag38511 Arabidops
44	1140	70.2	304	3	AAQ44512	Aag44512 Arabidops
45	1135	69.9	304	3	AAQ06649	Aag06649 Arabidops

ALIGNMENTS

RESULT 1  
ID AAW81018 standard; protein; 325 AA.

XX AAW81018;

DT 27-SEP-1999 (first entry)

XX

DE Fragment of cysteine synthase from soybean.

XX

KW Biosynthesis; biosynthetic pathway; lysine; threonine; methionine; cyteine; isoleucine; amino acid; homoserine kinase;

KW aspartic semialdehyde dehydrogenase; diaminopimelate decarboxylase;

KW cysteine synthase; cystathione beta-lyase; gene expression; screening; inhibition.

XX

OS Glycine max.

XX

PN WO9856935-A2.

XX

PD 17-DEC-1998.

XX

PF 11-JUN-1998; 98WO-US012073.

XX

PR 12-JUN-1997; 97US-0049406P.

PR 12-NOV-1997; 97US-0065385P.

XX

PA (DUPO ) DU PONT DE NEMOURS & CO E I.

XX

PI Falco SC, Allen SM, Thorpe CJ;

XX

DR WPI; 1999-080910/07.

DR N-PSDB; AAW99906.

XX

PT New isolated plant amino acid biosynthetic enzyme nucleic acids - which encode aspartic semialdehyde dehydrogenase, diaminopimelate decarboxylase, homoserine kinase, cysteine synthase and cystathionine beta-lyase.

PT

PS Claim 16; Page 58-59; 80pp; English.

XX

CC Organisation of the pathway leading to plant biosynthesis of lysine, threonine, methionine, cysteine and isoleucine suggests that over-expression or reduction of expression of genes encoding enzymes involved in that biosynthetic pathway could be used to alter the level of these amino acids in human food and animal feed. This may increase the nutritional quality of human food and animal feed by increasing the production and accumulation of specific free amino acids. The enzymes

CC include aspartic semialdehyde dehydrogenase, homoserine kinase,  
CC diaminopimelate decarboxylase, cysteine synthase and cystathione beta-  
CC lyase. The nucleic acids encoding these enzymes can be used for altering  
CC the level of expression of the enzymes and for evaluating compounds for  
CC their ability to inhibit the enzymes' activity  
XX

Query Match 100.0%; Score 1623; DB 2; Length 325;  
Best Local Similarity 100.0%; Pred. No. 1.1e-149;  
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVERSGIAKDVTELTIGKTPVLYLNKLDGCVARVAARVAALMEPCSSVKDRIGYSMTADA 60  
DB 1 MAVERSGIAKDVTELTIGKTPVLYLNKLDGCVARVAARVAALMEPCSSVKDRIGYSMTADA 60  
QY 61 EEKGLITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILLAFGAELV 120  
DB 61 EEKGLITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILLAFGAELV 120  
QY 121 LTDPAGKMGAVQKAEIILAKTNPAYILQGFENPANPKVHYETTGPETWKSDDKIDAFV 180  
DB 121 LTDPAGKMGAVQKAEIILAKTNPAYILQGFENPANPKVHYETTGPETWKSDDKIDAFV 180  
QY 181 SGIGTGTITGAGKYLKEQNPNIKLIGVEPVESPVLSGGKPGPHKIQIGAGFIPGVLEV 240  
DB 181 SGIGTGTITGAGKYLKEQNPNIKLIGVEPVESPVLSGGKPGPHKIQIGAGFIPGVLEV 240  
QY 241 NLIDEVVOISSDEALETAKLALKEGLFVGISGAAAAAFAQIAKRPENAGKLIIVAVFPS 300  
DB 241 NLIDEVVOISSDEALETAKLALKEGLFVGISGAAAAAFAQIAKRPENAGKLIIVAVFPS 300  
QY 301 FGERYLSVLPESVRRREASMTFEP 325  
DB 301 FGERYLSVLPESVRRREASMTFEP 325

RESULT 2  
AAR49830 standard; protein; 325 AA.  
ID AAR49830;  
AC AAR49830;  
XX  
DT 12-OCT-1994 (first entry)  
XX  
DE Cysteine synthase.  
XX  
KM Cysteine synthase; plant; expression; probe.  
XX  
OS Spinacia oleracea.  
XX  
PN JP06038770-A.  
XX  
PD 15-FEB-1994.  
XX  
PF 05-FEB-1992; 92JP-00020315.  
XX  
PR 05-FEB-1992; 92JP-00020315.  
XX  
PA (MITS) MITSUBISHI CORP.  
XX  
PA (MITU) MITSUBISHI KASEI CORP.  
XX  
DR WPI; 1994-094834/12.  
XX  
DR N-PSDB; AAQ44450.  
XX  
PT Novel gene coding cysteine synthase - used to increase the cysteine  
PT content of an agricultural plant.  
XX  
PS Claim 2; Page 4-6; 6pp; Japanese.  
XX  
CC The cysteine content in an agricultural product can be increased by  
CC expressing the cysteine synthase in a plant. Probes V822 and V812 used in  
CC the isolation of the gene are given in AAQ44483-84

XX  
SQ Sequence 325 AA;  
Query Match 88.2%; Score 1431; DB 2; Length 325;  
Best Local Similarity 86.7%; Pred. No. 6.6e-131;  
Matches 281; Conservative 24; Mismatches 19; Indels 0; Gaps 0;

QY 1 MAVERSGIAKDVTELTIGKTPVLYLNKLDGCVARVAARVAALMEPCSSVKDRIGYSMTADA 60  
DB 1 MAVERSGIAKDVTELTIGKTPVLYLNKLDGCVARVAARVAALMEPCSSVKDRIGYSMTADA 60  
QY 61 EEKGLITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILLAFGAELV 120  
DB 61 EEKGLITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILLAFGAELV 120  
QY 121 LTDPAGKMGAVQKAEIILAKTNPAYILQGFENPANPKVHYETTGPETWKSDDKIDAFV 180  
DB 121 LTDPAGKMGAVQKAEIILAKTNPAYILQGFENPANPKVHYETTGPETWKSDDKIDAFV 180  
QY 181 SGIGTGTITGAGKYLKEQNPNIKLIGVEPVESPVLSGGKPGPHKIQIGAGFIPGVLEV 240  
DB 181 SGIGTGTITGAGKYLKEQNPNIKLIGVEPVESPVLSGGKPGPHKIQIGAGFIPGVLEV 240  
QY 241 NLIDEVVOISSDEALETAKLALKEGLFVGISGAAAAAFAQIAKRPENAGKLIIVAVFPS 300  
DB 241 NLIDEVVOISSDEALETAKLALKEGLFVGISGAAAAAFAQIAKRPENAGKLIIVAVFPS 300  
QY 301 FGERYLSVLPESVRRREASMTFEP 324  
DB 301 FGERYLSVLPESVRRREASMTFEP 324

RESULT 3  
AAG29607 standard; protein; 322 AA.  
ID AAG29607;  
AC AAG29607;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 35257.  
XX  
KM Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.



PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 09-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 23-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.



PI Kloti A, Woessner J, Zayed A, Boyes D, Davis K, Hamilton C;  
PI Ascenzi R, Hoffman N;  
DR WPI; 2002-519673/55.  
XX  
PT Identifying compound as candidate herbicide, comprises contacting  
PT cysteine synthase with compound and detecting presence and/or absence of  
PT binding between compound and synthase.  
XX  
PS Claim 4; Page 25-26; 30pp; English.  
XX  
CC The invention relates to a method of identifying a compound as a  
CC candidate for a herbicide, comprising contacting cysteine synthase with  
CC the compound, and detecting the presence and/or absence of binding  
CC between the compound and cysteine synthase, where binding indicates that  
CC the compound is a candidate for a herbicide. Cysteine synthase is  
CC essential for plant growth. Specially, the inhibition of cysteine  
CC synthase gene expression in plant seedlings results in severe  
CC developmental abnormalities. Thus, cysteine synthase is used as a target  
CC for identification of herbicides. The present sequence is Arabidopsis  
CC thaliana cysteine synthase (EC 4.2.99.8) Atocly  
XX  
SQ Sequence 322 AA;

Query Match 85.3%; Score 1384; DB 5; Length 322;  
Best Local Similarity 83.4%; Pred. No. 2.5e-126;  
Matches 266; Conservative 30; Mismatches 23; Indels 0; Gaps 0;

QY 6 SGIAKDVTELTIGKTPLVYLNKLDGCVARVAAKLEIMPCSSVKDRIGYSMIADAEKGL 65  
DB 3 SRIKDVTELTIGKTPLVYLNKLDGCVARVAAKLEIMPCSSVKDRIGYSMIADAEKGL 62  
QY 66 ITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILAFGAEVLVTDPA 125  
DB 63 IKPGESVLIPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILAFGAEVLVTDPA 122  
QY 126 KGMKGAIVQKAEIILAKTPNAVYILQOFENPANPKVHYETTGPEIMWKSDDKIDAFVSGIGT 185  
DB 123 KGMKGAIVQKAEIILAKTPNAVYILQOFENPANPKVHYETTGPEIMWKSDDKIDAFVSGIGT 182  
QY 186 GGTITGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPTGVLEVNLDLDE 245  
DB 183 GGTITGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPTGVLEVNLDLDE 242  
QY 246 VVOIISDEAIETAKLALKEGLFVGISGAAAAAFAQIAKRPENAGKLIYAVFPSPFGERY 305  
DB 243 VVOIISDEAIETAKLALKEGLFVGISGAAAAAFAQIAKRPENAGKLIYAVFPSPFGERY 302  
QY 306 LSSVLFESVRRBAESMTFB 324  
DB 303 LSTVLFDAATKAEAMTFB 321

RESULT 6  
ABB94010  
ID ABB94010 standard; protein; 322 AA.  
XX  
AC ABB94010;  
XX  
DT 31-MAY-2002 (first entry)  
XX  
DE Herbicidally active polypeptide SEQ ID NO 3221.  
XX  
KM Herbicidal; plant; agriculture; herbicide.  
XX  
OS Arabidopsis thaliana.  
XX  
PN WO200210210-A2.  
XX  
PD 07-FEB-2002.  
XX  
PF 28-AUG-2001; 2001WO-EP009892.  
XX

PR 28-AUG-2001; 2001WO-EP009892.  
XX  
PA (FARB ) BAYER AG.  
XX  
PI Tietjen K, Weidler M;  
XX  
DR WPI; 2002-269010/31.  
XX

PT Identifying plant target proteins for herbicidally active compounds,  
PT comprising aligning and comparing nucleic acid or amino acid sequences  
PT from plant with nucleic acid or amino acid sequences from non-plant  
PT organisms.  
XX  
PS Claim 5; SEQ ID NO 3221; 261pp + Sequence listing; English.  
XX

CC The invention relates to identifying target proteins (ABB90790-ABB94016)  
CC for herbicidally active compounds, comprising aligning and comparing  
CC nucleic acid or amino acid sequences from plant with nucleic acid or  
CC amino acid sequences from non-plant organisms using suitable search  
CC parameters, where plant sequences having an E-value greater by a factor  
CC of 3 than the E-value of most similar non-plant sequences are selected.  
CC The polypeptides or nucleic acids encoding them are useful for  
CC identifying modulators. The identified modulators are useful as  
CC herbicides  
XX  
SQ Sequence 322 AA;

Query Match 85.0%; Score 1379; DB 5; Length 322;  
Best Local Similarity 83.1%; Pred. No. 7.8e-126;  
Matches 265; Conservative 30; Mismatches 24; Indels 0; Gaps 0;

QY 6 SGIAKDVTELTIGKTPLVYLNKLDGCVARVAAKLEIMPCSSVKDRIGYSMIADAEKGL 65  
DB 3 SRIKDVTELTIGKTPLVYLNKLDGCVARVAAKLEIMPCSSVKDRIGYSMIADAEKGL 62  
QY 66 ITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILAFGAEVLVTDPA 125  
DB 63 IKPGESVLIPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILAFGAEVLVTDPA 122  
QY 126 KGMKGAIVQKAEIILAKTPNAVYILQOFENPANPKVHYETTGPEIMWKSDDKIDAFVSGIGT 185  
DB 123 KGMKGAIVQKAEIILAKTPNAVYILQOFENPANPKVHYETTGPEIMWKSDDKIDAFVSGIGT 182  
QY 186 GGTITGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPTGVLEVNLDLDE 245  
DB 183 GGTITGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPTGVLEVNLDLDE 242  
QY 246 VVOIISDEAIETAKLALKEGLFVGISGAAAAAFAQIAKRPENAGKLIYAVFPSPFGERY 305  
DB 243 VVOIISDEAIETAKLALKEGLFVGISGAAAAAFAQIAKRPENAGKLIYAVFPSPFGERY 302  
QY 306 LSSVLFESVRRBAESMTFB 324  
DB 303 LSTVLFDAATKAEAMTFB 321

RESULT 7  
AAB14775  
ID AAB14775 standard; protein; 322 AA.  
XX  
AC AAB14775;  
XX  
DT 30-OCT-2002 (first entry)  
XX  
DE Arabidopsis thaliana cysteine synthase #2.  
XX  
KM Herbicide; cysteine synthase; plant growth; developmental abnormality;  
XX  
OS Arabidopsis thaliana.  
XX  
PN WO200246451-A2.  
XX



```

PD      13-JUN-2002.
XX
XX      25-OCT-2001; 2001WO-US050890.
XX
XX      26-OCT-2000; 2000US-00697225.
XX
XX      (PARA-) PARADIGM GENETICS INC.
XX
XX      Klotz A, Woessner J, Zayed A, Boyes D, Davis K, Hamilton C;
PI      Ascenzi R, Hoffman N;
XX
DR      WPI; 2002-519673/55.
XX
PT      Identifying compound as candidate herbicide, comprises contacting
PT      cysteine synthase with compound and detecting presence and/or absence of
PT      binding between compound and synthase.
XX
PS      Claim 4; Page 26; 30pp; English.
XX
CC      The invention relates to a method of identifying a compound as a
CC      candidate for a herbicide, comprising contacting cysteine synthase with
CC      the compound, and detecting the presence and/or absence of binding
CC      between the compound and cysteine synthase, where binding indicates that
CC      the compound is a candidate for a herbicide. Cysteine synthase is
CC      essential for plant growth. Specially, the inhibition of cysteine
CC      synthase gene expression in plant seedlings results in severe
CC      developmental abnormalities. Thus, cysteine synthase is used as a target
CC      for identification of herbicides. The present sequence is Arabidopsis
CC      thaliana cysteine synthase (EC 4.2.99.8) Atcys3A
XX
SQ      Sequence 322 AA;

Query Match          85.0%; Score 1379; DB 5; Length 322;
Best Local Similarity 83.1%; Pred. No. 7.8e-126;
Matches 265; Conservative 30; Mismatches 24; Indels 0; Gaps 0;

QY      6 SGIAKDVTELGKTPLVLYLNKLADGCVARVAAKLELMERPSSVXDRIGYSMLDAEKKGL 65
DQ      3 SRIAKDVTELGNTPLVLYLNNVABGCVGRAAKLEMEPCSSVXDRIGPSMTSDAEKKGL 62
QY      66 ITPGKSVLIEPTSGNTGIGLAFAAAARGYKLIITPMASMSLERIILLAFGAELVLTDP 125
DQ      63 IKPGESVLIEPTSGNTGVGLAFTAAAGKYKLIITPMASMSTERIILLAFGEVLVLTDP 122
QY      126 KGMKGAVOKAEILAKTPNAVYLQQFENPANPKVHYETTGPPEIKWGSDDKIDAPVSGIGT 185
DQ      123 KGMKGAIKAEEILAKTPNGVMLOQFENPANPKIHETTGPPEIKWGTSKIDGFVSGIGT 182
QY      186 GGTTTGAGKYLKEQNPNKILIGVEPYESPVLSGGKPGPHKIQGIGAGFIIPGVLEVNLIDE 245
DQ      183 GGTTTGAGKYLKEQNANVKLYGVEPYESAILSGKPGPHKIQIGAGFIIPSVLNVLDIDE 242
QY      246 VVOISSDEALETAKLLALKEGLFYGISGAAAAAFOJAKRPERNAGKLIIVAFPSPFGERY 305
DQ      243 VVOVSSDESIDMARQLAKEGLIYGISGABAATAIKLAQRPENAGKLFVALFPSPFGERY 302
QY      306 LSSVLPESVRREAESMTFE 324
DQ      303 LSTVLPDATRKCAEAAMTFE 321

RESULT 8
ID      ABM73727 standard; protein; 325 AA.
XX
XX      ABM73727;
AC
DT      17-OCT-2003 (first entry)
XX
XX      DNA clone originating in barley containing SNP sequence #137.
XX
XX      Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis.
XX
```

OS Hordeum vulgare.  
XX  
PN WO2003057877-A1.  
XX  
PD 17-JUL-2003.  
XX  
PF 16-DEC-2002; 2002WO-IB005403.  
XX  
PR 20-DEC-2001; 2001JP-00387059.  
PR 20-DEC-2001; 2001JP-00387131.  
PR 20-DEC-2001; 2001JP-00403289.  
PR 20-DEC-2001; 2001JP-00403300.  
PR 27-SEP-2002; 2002JP-00327515.  
XX  
PA (UTNI-) UNIV JAPAN OKAYAMA.  
XX  
PI Sato K, Takeda K, Kohara Y;  
XX  
DR WPI; 2003-587127/55.  
XX  
PT Single nucleotide polymorphism sites in barley varieties and DNA  
PT sequences containing them for analysis and identification of barley  
PT varieties and production of barley transformants with desired  
PT characteristics.  
XX  
PS Disclosure; SEQ ID XX; 284pp; Japanese.  
XX  
CC The present invention relates to oligonucleotide clones originating in  
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms  
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley  
CC varieties, identification of particular varieties and genotype-phenotype  
CC analysis, isolation of specific genes and creation of new varieties by  
CC transformation of barley varieties with them and production of new barley  
CC varieties with desired properties. The present sequence represents an  
CC oligonucleotide clone sequence featured in the specification. The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published-pct-sequences  
XX  
SQ Sequence 325 AA;

Query Match	84.3%	Score 1368	DB 7	Length 325
Best Local Similarity	83.0%	Pred. No. 9.4e-125		
Matches 263	Conservative 26	Mismatches 28	Indels 0	Gaps 0
QY	8	IADVTBLIGKTPLVYLNKLADGCVARVAAKLBLEPCCSVKORIGYMIADAEKGLIT	67	
Db	9	IADVTBLIGNTPLVYLNKRTDGCVRVAAKLBSMEPCSSVKORIGYSMTDAEKGFIY	68	
QY	68	PKGSVLI EPTSGNTGIGLAFMAAARGYKLIITMPASMSIERRIITLAFGAELVLTDPAKG	127	
Db	69	PKGSVLI EPTSGNTGIGLAFMAAARGYRLVLTMPASMSMERRIITLKAFGAELITDPLLG	128	
QY	128	MKGAVQKAEIILAKTPNAYILQOEFENPANPKVHYETTGPRIWKSDDKIDAFVSGIGTGG	187	
Db	129	MKGAVQKAEIILAKTPNSYILQOEFENANPKIHYETTGPRIWKTGDKIDGLVSGIGTGG	188	
QY	188	TITGAGKYLKEQNPNIKLIGVEPVESFVLSGKPGPHKIQIGAGFIPIGVLEVNLLDEVV	247	
Db	189	TITGSGKYLKEQNPNIKLYGVEPTESAVLNGKPGPHKIQIGAGFIPIGLVDVDIIDETV	248	
QY	248	QISSDEAIBTAKLTLALKEGLFVGISSGAAAAAFOIAKRPENAGKLIYAVFESFGERYLS	307	
Db	249	QVSSDESIEMAKSLALKEGLLVGISSGAAAAAIIKVAQRPENAGKLPYVVFPSFGERYLS	308	
QY	308	SVLFESVRRREAESMTFE	324	
Db	309	SVLFHSIKKEAESMTVE	325	

XX ABM73998;  
AC 17-OCT-2003 (first entry)  
DT DNA clone originating in barley containing SNP sequence #408.  
XX  
XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis.  
XX  
XX Hordeum vulgare.  
XX WO2003057877-A1.  
XX  
XX 17-JUL-2003.  
XX  
XX 16-DEC-2002; 2002WO-IB005403.  
XX  
XX 20-DEC-2001; 2001JP-00387059.  
XX 20-DEC-2001; 2001JP-00387131.  
XX 20-DEC-2001; 2001JP-00403299.  
XX 20-DEC-2001; 2001JP-00403300.  
XX 27-SEP-2002; 2002JP-00327515.  
XX  
XX (UTMI-) UNIV JAPAN OKAYAMA.  
XX  
XX Sato K, Takeda K, Kohara Y;  
XX WPI; 2003-587127/55.  
XX  
XX Single nucleotide polymorphism sites in barley varieties and DNA  
XX PT sequences containing them for analysis and identification of barley  
XX PT varieties and production of barley transformants with desired  
XX PT characteristics.  
XX  
XX PS Disclosure; SEQ ID XX; 284bp; Japanese.  
XX  
XX The present invention relates to oligonucleotide clones originating in  
XX CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms  
XX CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley  
XX CC varieties, identification of particular varieties and genotype-phenotype  
XX CC analysis, isolation of specific genes and creation of new varieties by  
XX CC transformation of barley varieties with them and production of new barley  
XX CC varieties with desired properties. The present sequence represents an  
XX CC oligonucleotide clone sequence featured in the specification. The  
XX CC sequence data for this patent did not form part of the printed  
XX CC specification, but was obtained in electronic format directly from WIPO  
XX CC at ftp.wipo.int/pub/published-pct-sequences  
XX  
XX SQ Sequence 418 AA;

Query Match 84.3%; Score 1368; DB 7; Length 418;  
Best Local Similarity 83.0%; Pred. No. 1.4e-124;  
Matches 263; Conservative 26; Mismatches 28; Indels 0; Gaps 0;  
OY 8 IAKDVTBLIGKTPPLVYLNKADGCVARVAAKLEIMPCSSVXDRIGYSMTADAEKGLIT 67  
Db 102 IARDVTBLIGNTPLVYLNKVTDCGVRVAALKESMBPCSSVXDRIGYSMTDAEEKGLITV 161  
OY 68 PKGSVLIERTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILAFGAELVLTDPKAG 127  
Db 162 PKGSVLIERTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILAFGAELVLTDPKAG 221  
OY 128 MKGAVQKAEELAKTPNAYILQGFENANPKVHYETTGPEIWKSGSDGKIDAFVSGIGTGG 187  
Db 222 MKGAVQKAEELAKTPNAYILQGFENANPKVHYETTGPEIWKSGSDGKIDAFVSGIGTGG 281  
OY 188 TITGAGKYLKQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIIPGVLBNLDEYV 247  
Db 282 TITGAGKYLKQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIIPGVLBNLDEYV 341  
OY 248 QISSDEALITAKLALKEGLFVGISGGAALAAFOIAKRPENAGKLIIVAFPSFGERYLS 307  
Db 342 QVSSDESLIMAKSLALKEGLFVGISGGAALAAFOIAKRPENAGKLIIVAFPSFGERYLS 401

OY 308 SVLFESVRRRAESMTFR 324  
Db 402 SVLFHSIKKBAESMVR 418  
RESULT 10  
ABB93998  
ID ABB93998 standard; protein; 305 AA.  
XX  
XX ABB93998;  
XX  
XX 31-MAY-2002 (first entry)  
XX  
XX Herbicidally active polypeptide SEQ ID NO 3209.  
XX DE Herbicidally active polypeptide SEQ ID NO 3209.  
XX  
XX Herbicidal; plant; agriculture; herbicide.  
XX  
XX Arabidopsis thaliana.  
XX OS  
XX WO200210210-A2.  
XX  
XX 07-FEB-2002.  
XX  
XX 28-AUG-2001; 2001WO-EP009892.  
XX  
XX 28-AUG-2001; 2001WO-EP009892.  
XX  
XX (FARB ) BAYER AG.  
XX  
XX Tietjen K, Weidner M;  
XX  
XX WPI; 2002-269010/31.  
XX  
XX PT Identifying plant target proteins for herbicidally active compounds,  
XX PT comprising aligning and comparing nucleic acid or amino acid sequences  
XX PT from plant with nucleic acid or amino acid sequences from non-plant  
XX PT organisms.  
XX  
XX PS Claim 5; SEQ ID NO 3209; 261bp + Sequence Listing; English.  
XX  
XX The invention relates to identifying target proteins (ABB90790-ABB94016)  
XX CC for herbicidally active compounds, comprising aligning and comparing  
XX CC nucleic acid or amino acid sequences from plant with nucleic acid or  
XX CC amino acid sequences from non-plant organisms using suitable search  
XX CC parameters, where plant sequences having an E-value greater by a factor  
XX CC of 3 than the E-value of most similar non-plant sequences are selected.  
XX CC The polypeptides or nucleic acids encoding them are useful for  
XX CC identifying modulators. The identified modulators are useful as  
XX CC herbicides  
XX  
XX SQ Sequence 305 AA;

Query Match 79.9%; Score 1296; DB 5; Length 305;  
Best Local Similarity 78.2%; Pred. No. 9.2e-118;  
Matches 254; Conservative 28; Mismatches 23; Indels 20; Gaps 1;  
OY 1 MAVERSGIAKDVTELIGKTPPLVYLNKADGCVARVAAKLEIMPCSSVXDRIGYSMTADA 60  
Db 1 MASVAPKIAKDVTELIGKTPPLVYLNKADGCVARVAAKLEIMPCSSVXDRIGYSMTADA 60  
OY 61 BEKGLITPKGSVLIERTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILAFGAELV 120  
Db 61 EAKGLIKPGBSVLIERTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILAFGAELV 120  
OY 121 LTPDAKMGKAVQKAEELAKTPNAYILQGFENANPKVHYETTGPEIWKSGSDGKIDAFV 180  
Db 121 LTPDAKMGKAVQKAEELAKTPNAYILQGFENANPKVHYETTGPEIWKSGSDGKIDAFV 170  
OY 181 SGIGTGTITGAGKYLKQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIIPGVLIDV 240  
Db 171 -----GAGKYLKQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIIPGVLIDV 220

QY 241 NLIDEVVOISSDEALETAKLALKEGLFVGISGAAAAAFOIAKRPENAGKLIIVAFPS 300  
Db 221 DLIDEVVOSSSESIDMARIAREGLLVGISSGAAATTAIKAKRPENAGKLIIVAFPS 280  
QY 301 FGERYLSVLPESVRRREASMTPEP 325  
Db 281 FGERYLSVLPFDAARKEALMTPEP 305

RESULT 11

AAR63756  
ID AAR63756 standard; protein; 383 AA.

AC AAR63756;

DT 13-JUN-1995 (first entry)

DE Spinach cysteine synthase.

KM Cysteine; spinach; spinach oleracea; pUC19; M13mpl8; plant; nutrition;  
KW feed value; pharmaceutical; food additive; cosmetic; E.coli.

OS Spinacia oleracea.

PN JP06245773-A.

PD 06-SEP-1994.

PF 26-FEB-1993; 93JP-00038527.

PR 26-FEB-1993; 93JP-00038527.

PA (MITS) MITSUBISHI CORP.

PA (MITU) MITSUBISHI KASEI CORP.

DR WPI; 1994-321282/40.

DR N-PSDB; AAQ74413.

PT Gene encoding a cysteine synthesis enzyme - useful for elevating cysteine  
PT content in plant bodies increasing their nutritional value.

PS Claim 2; Page 4-6; 6pp; Japanese.

CC The amino acid sequence of the novel cysteine synthase enzyme from  
CC spinach, Spinacia oleracea. The probes (AAQ74414-5) were used to obtain a  
CC clone from a cDNA library derived from spinach seedling leaves RNA. The  
CC 1.5 kb insert was ligated into the cloning vectors pUC19 and M13 mp18.

CC The gene, 1483 bp, encodes a protein of 383 a.a. The cysteine synthase  
CC gene is expressed in plants to elevate the cysteine content in the plant.  
CC The nutritional and feed value of the plant are expected to be enriched.

CC Cysteine, one of the S-containing amino acids, can be used as materials  
CC for various pharmaceuticals, food additives or cosmetics. Production of  
CC the protein is a useful step in the production of cysteine

XX Sequence 383 AA;

XX

Query Match 78.3%; Score 1271; DB 2; Length 383;

Best Local Similarity 73.4%; Pred. No. 3.6e-115;

Matches 234; Conservative 52; Mismatches 33; Indels 0; Gaps 0;

QY 3 VERSGIAXDVTTELIGKTPVLYLNKLAGCVARVAAXEIMBPCCSVKDRIGYSMTADAE 62

Db 62 IEGNTIAEDVSQIGKTPMYLNNVSKGSVANIAXLESMBPCSVKDRIGYSMTADAE 121

QY 63 KGLITPGKSVLIEPTSGNTGIGLAFMAAARGVYKLTITMPASMSLERITLLAFGAELVLT 122

Db 122 KGVITPGKTVLIEPTSGNTGIGLAFIAAARGVYKLTITMPASMSMERITLLAFGAELVLT 181

QY 123 DPAKMGKAVOKABEILATPNAYILQOFENPANPKVHETTGPEIWKSGDKIDAFVSG 182

Db 182 DPAKMGKAVOKABEILATPNAYILQOFENPANPKVHETTGPEIWKSGDKIDAFVSG 241

QY 183 IGTGTTTGAGKYLKEQNPNIKLIGVEPVASPVLSGKPGPHKIQIGAGPIPGVLEVN 242

Db 242 IGTGTTTGAGKYLKEQNPNIKLIGVEPVASPVLSGKPGPHKIQIGAGPIPGVLEVN 301  
QY 243 LDEVVOISSDEALETAKLALKEGLFVGISGAAAAAFOIAKRPENAGKLIIVAFPSFG 302  
Db 302 MDEVIVSSSEBAVEMAKQIAMEGGLVGISSGAAAAAFAVRIGKRPENAGKLIIVAFPSFG 361  
QY 303 ERYLSSVLPESVRRREASMTPEP 321  
Db 362 ERYLSSVLPESVRRREASMTPEP 380

RESULT 12

AAQ09478  
ID AAQ09478 standard; protein; 392 AA.

AC AAQ09478;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 7430.

KM Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-00301439.

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.

PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.

PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137502P.

PR 07-JUN-1999; 99US-0137724P.



PR	08-JUN-1999;	99US-0138094P.	PR	09-AUG-1999;	99US-0147493P.
PR	10-JUN-1999;	99US-0138540P.	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	99US-0138847P.	PR	10-AUG-1999;	99US-0148171P.
PR	14-JUN-1999;	99US-0139119P.	PR	11-AUG-1999;	99US-0148319P.
PR	16-JUN-1999;	99US-0139452P.	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	99US-0139453P.	PR	13-AUG-1999;	99US-0148555P.
PR	17-JUN-1999;	99US-0139492P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139454P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139455P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139456P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139457P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	22-JUN-1999;	99US-0139899P.	PR	30-AUG-1999;	99US-0151303P.
PR	23-JUN-1999;	99US-0140353P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0152363P.
PR	28-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.
PR	29-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144325P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	21-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161992P.
PR	04-AUG-1999;	99US-0147204P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147302P.	PR	29-OCT-1999;	99US-0162142P.
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			

Query Match 77.7%; Score 1261; DB 3; Length 392;  
 Best Local Similarity 74.9%; Pred. No. 3.5e-114;  
 Matches 239; Conservative 41; Mismatches 39; Indels 0; Gaps 0;

QY 3 VERSGIADVTTELIGKTPVLYLNKLADGCVARVAAKLEIMEPCSSVKDRIGYSMLADAE 62  
 Db 70 VEGNLADNAQILIGKTPVLYLNKLADGCVARVAAKLEIMEPCSSVKDRIGYSMLADAE 129  
 QY 63 KGLITPGKSVLIEPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILAFGAELVLT 122  
 Db 130 KGLITPGKSVLIEPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILAFGAELVLT 189  
 QY 123 DPAKGMKGAVOKABEILAKTPNAYILQOFENPANPKVHYETTGPETIWKSGDGIKIDAFVSG 182  
 Db 190 BPAKGMTGAIQKABEILAKTPNAYILQOFENPANPKVHYETTGPETIWKSGDGIKIDAFVSG 249  
 QY 183 IGTGCTITGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGIGAGFIPGVLEVNL 242  
 Db 250 IGTGCTITGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGIGAGFIPGVLEVNL 309  
 QY 243 LDEVVQISSDEAIFETAKLLALKEGLFVGISSGAAAAAFOIAKRPENAGKELIIVFPSPG 302  
 Db 310 VDEYIAISSERAIETSKQALQEGLLVGISSGAAAAAFOIAKRPENAGKELIIVFPSPG 369  
 QY 303 ERYLSSVLFESVRRREASM 321  
 Db 370 ERYLSTOLFQISIRECEQM 388

RESULT 13

ABB94007  
 ID ABB94007 standard; protein; 442 AA.

AC ABB94007;

DT 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 3218.

KN Herbicidal; plant; agriculture; herbicide.

OS Arabidopsis thaliana.

PN WO200210210-A2.

PD 07-FEB-2002.

PF 28-AUG-2001; 2001WO-EP009892.

PR 28-AUG-2001; 2001WO-EP009892.

PA (FARB ) BAYER AG.

PI Tietjen K, Weidler M;

DR WPI; 2002-269010/31.

PT Identifying plant target proteins for herbicidally active compounds,

PT comprising aligning and comparing nucleic acid or amino acid sequences

PT from plant with nucleic acid or amino acid sequences from non-plant

PT organisms.

PS Claim 5; SEQ ID NO 3218; 261PP + Sequence listing; English.

XX The invention relates to identifying target proteins (ABB90790-ABB94016)

CC for herbicidally active compounds, comprising aligning and comparing

CC nucleic acid or amino acid sequences from plant with nucleic acid or

CC amino acid sequences from non-plant organisms using suitable search

CC parameters, where plant sequences having an E-value greater by a factor

CC of 3 than the E-value of most similar non-plant sequences are selected.

CC The polypeptides or nucleic acids encoding them are useful for

CC identifying modulators. The identified modulators are useful as

Best Local Similarity 71.3%; Pred. No. 1.6e-113;  
 Matches 241; Conservative 43; Mismatches 36; Indels 18; Gaps 2;

QY 2 AVERS-----GIADVTTELIGKTPVLYLNKLADGCVARVAAKLEIMEPCSSVKDRIGYS 55  
 Db 101 AVKRETPDGINLADNVSQLIGKTPVLYLNKLADGCVARVAAKLEIMEPCSSVKDRIGYS 160  
 QY 56 MIADAEKGLITPGKSVLIEPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILAF 115  
 Db 161 MYTDAEQKGFISPGKSVLIEPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILAF 220  
 QY 116 GAELVLTDPAGKMGAVOKABEILAKTPNAYILQOFENPANPKVHYETTGPETIWKSGDK 175  
 Db 221 GAELVLTDPAGKMGAVOKABEILAKTPNAYILQOFENPANPKVHYETTGPETIWKSGDK 280  
 QY 176 IDAFVSGIGTGTITGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGIGAGFIP 235  
 Db 281 VDFVAGIGTGTITGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGIGAGFIP 340  
 QY 236 GVLBNVLLDEVVQISSDEAIFETAKLLALKEGLFVGISSGAAAAAFOIAKRPENAGKLI 295  
 Db 341 KMLDQKIMDEVIAISSERAIETSKQALQEGLLVGISSGAAAAAFOIAKRPENAGKLI 400  
 QY 296 -----AVFSPFGERYLSVLFESVRRREASM 321  
 Db 401 VTVNESMNAVFQVFPSPFGERYLSVLFESVRRREASM 438

RESULT 14

ABB94013  
 ID ABB94013 standard; protein; 315 AA.

AC ABB94013;

DT 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 3224.

KN Herbicidal; plant; agriculture; herbicide.

OS Arabidopsis thaliana.

PN WO200210210-A2.

PD 07-FEB-2002.

PF 28-AUG-2001; 2001WO-EP009892.

PR 28-AUG-2001; 2001WO-EP009892.

PA (FARB ) BAYER AG.

PI Tietjen K, Weidler M;

DR WPI; 2002-269010/31.

PT Identifying plant target proteins for herbicidally active compounds,

PT comprising aligning and comparing nucleic acid or amino acid sequences

PT from plant with nucleic acid or amino acid sequences from non-plant

PT organisms.

PS Claim 5; SEQ ID NO 3224; 261PP + Sequence listing; English.

XX The invention relates to identifying target proteins (ABB90790-ABB94016)

CC for herbicidally active compounds, comprising aligning and comparing

CC nucleic acid or amino acid sequences from plant with nucleic acid or

CC amino acid sequences from non-plant organisms using suitable search

CC parameters, where plant sequences having an E-value greater by a factor

CC of 3 than the E-value of most similar non-plant sequences are selected.

CC The polypeptides or nucleic acids encoding them are useful for

CC identifying modulators. The identified modulators are useful as

Seq Sequence 315 AA;

Query Match 77.1%; Score 1252; DB 5; Length 315;  
Best Local Similarity 80.3%; Pred. No. 1.9e-113;  
Matches 245; Conservative 28; Mismatches 30; Indels 2; Gaps 1;

QY 6 SGIADVTTELIGKTPLVYLNKLADGCYARVAAKLELMEPCSSVKDRIGYSMTADAEKGL 65  
DB 3 SRIADVTTELIGNTLLVYLNVAEGCVRAAKLELMEPCSSVKDRIGYSMTADAEKGL 62  
QY 66 ITPGKSVLIBPTSGNTGIGLAFMAARGYKLIITMPASMSLERRIILAFGAELVLTDP 125  
DB 63 IKPGESVLIBPTSGNTGIGLAFMAARGYKLIITMPASMSLERRIILAFGAELVLTDP 122  
QY 126 KMKGAVQKAEIILAKTPNAVYILQGFENPANKVHYETTGPEIWKSGDKIDAFVSGIGT 185  
DB 123 KMKGALAKAEIILAKTPNGYMLQGFENPANKIHYETTGPEIWKSGDKIDAFVSGIGT 182  
QY 186 GGTITGAGKYLKEQNPNKILGVEPVESPVLSGKPGPHKIQIGAGFIPSVLNVLDL 245  
DB 183 G--ITGASILKRNRTANVKLYGVEPVESAAILIGKPGPHKIQIGAGFIPSVLNVLDL 240  
QY 246 VVQISSDEAIEITAKLLALKEGLFVGISGAAAAAFAQIAKRPENAGKLIYAVPSPFGERY 305  
DB 241 VVQVSSDESIDMARQALKEGLFVGISGAAAAAFAQIAKRPENAGKLIYAVPSPFGERY 300  
QY 306 ESSVL 310  
DB 301 ESTVL 305

RESULT 15

ABB94001

ID ABB94001 standard; protein; 315 AA.

AC ABB94001;

DT 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 3212.

KM Herbicidal; plant; agriculture; herbicide.

OS Arabidopsis thaliana.

PN WO200210210-A2.

PD 07-FEB-2002.

PF 28-AUG-2001; 2001WO-BP009892.

PR 28-AUG-2001; 2001WO-BP009892.

PA (FARB ) BAYER AG.

PI Tietjen K, Weidler M;

DR WPI; 2002-269010/31.

PT Identifying plant target proteins for herbicidally active compounds,  
comprising aligning and comparing nucleic acid or amino acid sequences  
from plant with nucleic acid or amino acid sequences from non-plant  
organisms.

PS Claim 5; SEQ ID NO 3212; 261pp + Sequence listing; English.

CC The invention relates to identifying target proteins (ABB90790-ABB94016)

CC for herbicidally active compounds, comprising aligning and comparing

CC nucleic acid or amino acid sequences from plant with nucleic acid or

CC amino acid sequences from non-plant organisms using suitable search

CC parameters, where plant sequences having an E-value greater by a factor

CC of 3 than the E-value of most similar non-plant sequences are selected.

CC identifying modulators. The identified modulators are useful as  
herbicides  
XX  
SQ Sequence 315 AA;

Query Match 77.1%; Score 1252; DB 5; Length 315;  
Best Local Similarity 80.3%; Pred. No. 1.9e-113;  
Matches 245; Conservative 28; Mismatches 30; Indels 2; Gaps 1;

QY 6 SGIADVTTELIGKTPLVYLNKLADGCYARVAAKLELMEPCSSVKDRIGYSMTADAEKGL 65  
DB 3 SRIADVTTELIGNTLLVYLNVAEGCVRAAKLELMEPCSSVKDRIGYSMTADAEKGL 62  
QY 66 ITPGKSVLIBPTSGNTGIGLAFMAARGYKLIITMPASMSLERRIILAFGAELVLTDP 125  
DB 63 IKPGESVLIBPTSGNTGIGLAFMAARGYKLIITMPASMSLERRIILAFGAELVLTDP 122  
QY 126 KMKGAVQKAEIILAKTPNAVYILQGFENPANKVHYETTGPEIWKSGDKIDAFVSGIGT 185  
DB 123 KMKGALAKAEIILAKTPNGYMLQGFENPANKIHYETTGPEIWKSGDKIDAFVSGIGT 182  
QY 186 GGTITGAGKYLKEQNPNKILGVEPVESPVLSGKPGPHKIQIGAGFIPSVLNVLDL 245  
DB 183 G--ITGASILKRNRTANVKLYGVEPVESAAILIGKPGPHKIQIGAGFIPSVLNVLDL 240  
QY 246 VVQISSDEAIEITAKLLALKEGLFVGISGAAAAAFAQIAKRPENAGKLIYAVPSPFGERY 305  
DB 241 VVQVSSDESIDMARQALKEGLFVGISGAAAAAFAQIAKRPENAGKLIYAVPSPFGERY 300  
QY 306 ESSVL 310  
DB 301 ESTVL 305

Search completed: May 18, 2004, 13:07:06  
Job time : 63 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2004, 13:04:09 ; Search time 21 Seconds

(without alignments)  
1488.678 Million cell updates/sec

Title: US-09-931-457a-31

Perfect score: 1623

Sequence: 1 MAVERSGIAKDVTETLIGKTP.....LGSVLFESVRRRAESMTFEP 325

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_78:\*

1: pirl:\*

2: pirl2:\*

3: pirl3:\*

4: pirl4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1453	89.5	325	1	S46438	cysteine synthase
2	1435	88.4	325	2	T07001	cysteine synthase
3	1431	88.2	325	2	S35094	cysteine synthase
4	1384	85.3	322	2	A71412	cysteine synthase
5	1378	84.9	325	1	S52738	cysteine synthase
6	1364	84.0	325	1	JS0762	cysteine synthase
7	1314	81.0	386	2	T07002	cysteine synthase
8	1271	78.3	387	2	T52650	cysteine synthase
9	1270	78.3	352	2	T07962	probable cysteine
10	1261	77.7	383	2	S29733	cysteine synthase
11	1261	77.7	392	2	A84870	cysteine synthase
12	1255	77.3	442	2	T47800	cysteine synthase
13	1252	77.1	315	2	S48694	cysteine synthase
14	1249	77.0	374	1	A43407	cysteine synthase
15	1229	75.7	324	2	S49586	cysteine synthase
16	1179	72.6	324	2	T52609	cysteine synthase
17	1170	72.1	390	2	T09000	cysteine synthase
18	1163	71.7	392	2	S48695	cysteine synthase
19	1047	64.5	368	2	A55450	cysteine synthase
20	1017	62.7	368	2	T47936	cysteine synthase
21	1004	61.9	319	2	AB2121	cysteine synthase
22	989	60.9	320	2	AH2374	cysteine synthase
23	971	59.8	331	2	S77347	cysteine synthase
24	968	59.6	344	2	T19367	cysteine synthase
25	949.5	58.5	310	2	G70660	cysteine synthase
26	947	58.3	322	2	AH2614	cysteine synthase
27	947	58.3	322	2	G97396	cysteine synthase
28	940.5	57.9	310	2	T44912	cysteine synthase
29	925	57.0	337	2	T23591	cysteine synthase

30	916	56.4	336	2	AH3264	cysteine synthase
31	912	56.2	310	2	H81161	cysteine synthase
32	896.5	55.2	309	2	E97175	cysteine synthase
33	882	54.3	322	2	C82258	cysteine synthase
34	878	54.1	337	2	C89009	cysteine synthase
35	877	54.0	308	2	T44614	cysteine synthase
36	869.5	53.6	329	1	S55321	cysteine synthase
37	865.5	53.3	323	2	AD0810	cysteine synthase
38	862	53.1	322	2	AI0363	cysteine synthase
39	860.5	53.0	328	2	G87698	cysteine synthase
40	859	52.9	308	2	S66103	cysteine synthase
41	857.5	52.8	323	2	SYBCAC	cysteine synthase
42	857.5	52.8	323	2	A85884	cysteine synthase
43	857.5	52.8	323	2	F91039	cysteine synthase
44	856	52.7	324	2	E83306	cysteine synthase
45	851.5	52.5	323	1	SYBAC	cysteine synthase

## ALIGNMENTS

## RESULT 1

S46438

cysteine synthase (BC 4.2.99.8) - watermelon

C/Species: Citrullus lanatus (watermelon)

C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000

C/Accession: S46438

R/Noji, M.; Murakoshi, I.; Saito, K.

Mol. Gen. Genet. 244, 57-66, 1994

A/Title: Molecular cloning of a cysteine synthase cDNA from Citrullus vulgaris (waterme

A/Reference number: S46438; MUID:94316193; PMID:8041362

A/Accession: S46438

A/Molecule type: mRNA

A/Residues: 1-325 <NO>

A/Cross-references: EMBL:D28777; MID:G466530; PIDN:BA05965.1; PID:G540497

A/Note: the source is designated as Citrullus vulgaris

A/Note: additional initiators could be Val-32 (GTG) or Met-41

C/Superfamily: threonine dehydratase

C/Keywords: carbon-oxygen lyase; cysteine biosynthesis; phosphoprotein; pyridoxal phosph

F;49/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 89.5%; Score 1453; DB 1; Length 325;

Best Local Similarity 87.1%; Pred. No. 1.6e-99;

Matches 283; Conservative 24; Mismatches 18; Indels 0; Gaps 0;

QY	1	MAVERSGIAKDVTETLIGKTPLVYLNKLDGCVARVAAXLILMBPOSSVKDRIGYMTADA	60
DB	1	MADAKSTIAKDVTETLIGKTPLVYLNKLDGCVARVAAXLILMBPOSSVKDRIGYMTADA	60
QY	61	BEKGLTFPGKSVLIBPTSGNTGIGLAFMAAARGYKLIITMPASMSLERIILAFGAEIV	120
DB	61	ENKGLTFPGKSVLIBPTSGNTGIGLAFMAAARGYKLIITMPASMSLERIILAFGAEIV	120
QY	121	LTPPAKMGKAVOKAEIILAKTPNAYILQGFENPANPKVHYETTGPEIWKSGDKIDAFV	180
DB	121	LTPPAKMGKAVOKAEIILAKTPNAYILQGFENPANPKVHYETTGPEIWKSGDKIDAFV	180
QY	181	SGIGTGTTGAGKYLKEQNPNIKLIGVPEVPSVPLSGKPGPHKIQIGAGFIPGVLEV	240
DB	181	SGIGTGTTGAGKYLKEQNPNIKLIGVPEVPSVPLSGKPGPHKIQIGAGFIPGVLEV	240
QY	241	NLDEVVQISSDEAIFTAKLALKEGLFVGIISSGAAAAAFAQIAKRPENAGKLIIVAFPS	300
DB	241	NLDEVVQISSDEAIFTAKLALKEGLFVGIISSGAAAAAFAQIAKRPENAGKLIIVAFPS	300
QY	301	FGERYLSSVLFESVRRRAESMTFEP 325	
DB	301	FGERYLSSVLFESVRRRAESMTFEP 325	

## RESULT 2

T07001

cysteine synthase (BC 4.2.99.8), cytosolic - potato



QY 66 ITPGKSVLIEPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILLAFGAEVLTDPA 125  
DB 63 IKPGESVLIPTSGNTGVGLAFAAARGYKLIITMPASMSLERRIILLAFGAEVLTDPA 122  
QY 126 KMKGAVOKAEELIAKTPNAVYIIOFENPANKVHYETGPEIWKSGDKIDAFVSGIGT 185  
DB 123 KMKGALAKAEELIAKTPNGVYLQOFENPANKIHYETGPEIWKGTGKIDGFWSGIGT 182  
QY 186 GGTITGAGKYLKEQNPNKILIGVEPVSSPVLISGKPGPHKIIGIGAGFIIPGVLBNVLTDE 245  
DB 183 GGTITGAGKYLKEQNANVKLYGVEPVESAILISGKPGPHKIIGIGAGFIIPSVLNDLIDE 242  
QY 246 VVOISSDEAIEIAKLALKEGLFVGISGAAAFQIAKRPENAGKLIIVAVFPFSGERY 305  
DB 243 VVOISSDESIDMARQIALKEGLVGISGAAAFQIAKRPENAGKLIIVAVFPFSGERY 302  
QY 306 LSSVLPFESVRRASMTPE 324  
DB 303 LSTVLPDATRKAEAMTPE 321

## RESULT 5

cysteine synthase (EC 4.2.99.8) precursor - maize  
S52738  
C:Species: Zea mays (maize)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: S52738  
R:Brander, K.A.; Owttrim, G.W.; Brumold, C.  
submitted to the EMBL Data Library, March 1995  
A:Description: Isolation of a putative plastidic isoform of cysteine synthase from maize  
A:Reference number: S52738  
A:Accession: S52738  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-325 <BRA>  
A:Cross-references: EMBL:X85803; NID:g758352; PIDN:CAA59798.1; PID:g758353  
C:Superfamily: threonine dehydratase  
C:Keywords: carbon-oxygen lyase; phosphoprotein; pyridoxal phosphate  
F:49/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 84.9%; Score 1378; DB 1; Length 325;  
Best Local Similarity 84.3%; Pred. No. 5.3e-94;  
Matches 268; Conservative 22; Mismatches 28; Indels 0; Gaps 0;

QY 8 IAKDVTELIIGTPVLYLNKLDGCVARVAALKEMLBPCSSVKDRIGYSMTADAEKGLIT 67  
DB 8 IAKDVTELIIGTPVLYLNKLDGCVGRSRAKLESMBCSSVKDRIGYSMTADAEKGLIT 67  
QY 68 PKGSVLIPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILLAFGAEVLTDPAKG 127  
DB 68 PKGSVLIPTSGNTGIGLAFAAARGYKLIITMPASMSMERRIILKAFGAEVLTDPLLG 127  
QY 128 MKGAVOKAEELIAKTPNAVYIIOFENPANKVHYETGPEIWKSGDKIDAFVSGIGTGG 187  
DB 128 MKGAVOKAEELIAKTPNSYIIOFENPANKIHYETGPEIWKATAGIDELVSGIGTGG 187  
QY 188 TITGAGKYLKEQNPNKILIGVEPVSSPVLISGKPGPHKIIGIGAGFIIPGVLBNVLTDE 247  
DB 188 TITGAGKYLKEQNPNKILIGVEPVESAVLNGKPGPHKIIGIGAGFIIPGVLBNVLTDE 247  
QY 248 QISSDEAIEIAKLALKEGLFVGISGAAAFQIAKRPENAGKLIIVAVFPFSGERYLS 307  
DB 248 QVSSDEAIEIAKLALKEGLLVGISSGAAAFQIAKRPENAGKLIIVAVFPFSGERYLS 307  
QY 308 SVLPFESVRRASMTPE 325  
DB 308 SVLPFESVRRASMTPE 325

RESULT 6  
JS0762  
cysteine synthase (EC 4.2.99.8) precursor - wheat

N:Alternate names: O-acetylserine (thiol)-lyase  
C:Species: Triticum aestivum (common wheat)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000  
C:Accession: JS0762  
R:Youssefian, S.; Nakamura, M.; Sano, H.  
submitted to JIPID, September 1992  
A:Reference number: JS0762  
A:Accession: JS0762  
A:Molecule type: mRNA  
A:Residues: 1-325 <YOU>  
A:Cross-references: DDBJ:D13153; NID:g218334; PIDN:BA02438.1; PID:g218335  
C:Superfamily: threonine dehydratase  
C:Keywords: carbon-oxygen lyase; cysteine biosynthesis; phosphoprotein; plastid; pyridic  
F:1-37/Domain: transit peptide (plastid) #status predicted <TMP>  
F:38-325/Product: cysteine synthase #status predicted <MAT>  
F:226/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 84.0%; Score 1364; DB 1; Length 325;  
Best Local Similarity 82.3%; Pred. No. 5.7e-93;  
Matches 261; Conservative 27; Mismatches 29; Indels 0; Gaps 0;

QY 8 IAKDVTELIIGTPVLYLNKLDGCVARVAALKEMLBPCSSVKDRIGYSMTADAEKGLIT 67  
DB 9 IAKDVTELIIGTPVLYLNKLDGCVGRVAALKEMLBPCSSVKDRIGYSMTADAEKGLIT 68  
QY 68 PKGSVLIPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILLAFGAEVLTDPAKG 127  
DB 69 PKGSVLIPTSGNTGIGLAFAAARGYKLIITMPASMSMERRIILKAFGAEVLTDPLLG 128  
QY 128 MKGAVOKAEELIAKTPNAVYIIOFENPANKVHYETGPEIWKSGDKIDAFVSGIGTGG 187  
DB 129 MKGAVOKAEELIAKTPNSYIIOFENPANKIHYETGPEIWKGTGKIDGVLGIGTGG 188  
QY 188 TITGAGKYLKEQNPNKILIGVEPVSSPVLISGKPGPHKIIGIGAGFIIPGVLBNVLTDE 247  
DB 189 TITGAGKYLKEQNPNKILIGVEPVESAVLNGKPGPHKIIGIGAGFIIPGVLBNVLTDE 248  
QY 248 QISSDEAIEIAKLALKEGLFVGISGAAAFQIAKRPENAGKLIIVAVFPFSGERYLS 307  
DB 249 QVSSDEAIEIAKLALKEGLLVGISSGAAAFQIAKRPENAGKLIIVAVFPFSGERYLS 308  
QY 308 SVLPFESVRRASMTPE 324  
DB 309 SVLPFESVRRASMTPE 325

## RESULT 7

T07002  
cysteine synthase (EC 4.2.99.8) precursor, chloroplast - potato  
C:Species: Solanum tuberosum (potato)  
C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 22-Jun-1999  
C:Accession: T07002  
R:Hesse, H.  
submitted to the EMBL Data Library, January 1998  
A:Reference number: Z15854  
A:Accession: T07002  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-386 <HES>  
A:Cross-references: EMBL:AF044173; NID:g3290021; PIDN:AAC25636.1; PID:g3290022  
A:Experimental source: cv. Beroлина  
C:Genetics:  
A:Genome: nuclear  
C:Function:  
A:Description: forms cysteine from O-acetyl-serine and hydrogen sulfide and releases cc  
C:Superfamily: threonine dehydratase  
C:Keywords: carbon-oxygen lyase; chloroplast; cysteine biosynthesis; phosphoprotein; py  
F:110/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 81.0%; Score 1314; DB 2; Length 386;  
Best Local Similarity 78.1%; Pred. No. 3.4e-89;  
Matches 249; Conservative 38; Mismatches 32; Indels 0; Gaps 0;



QY 3 VERSGIADVTTELIGTPTLVYLNKLADGCVARVAAKLELMEPCSSVKRIGYSMIADAE 62  
Db 64 IEGNIAEDVTOLIGNTPMVYLNLTIAKGCVANIAAKLEIMEPCSSVKRIGYSMTADAE 123  
QY 63 KGLITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERIILLAFGAEVLV 122  
Db 124 KGLISPGKTVLVEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERIILLAFGAEVLV 183  
QY 123 DPAKMGKAVOKAEIILAKTPNAYILOQFENPANPKVHYETGPEIWKSDGKIDAFVSG 182  
Db 184 DPAKMGKAVOKAEIILAKTPNAYILOQFENPANPKVHYETGPEIWKSDGKIDAFVSG 243  
QY 183 IGTGTTTGAGKYLEKQNPNIKIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEVN 242  
Db 244 IGTGTTTGAGKYLEKQNPNIKIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEVN 303  
QY 243 LDEVVOISSDEAIFETAKLALKEGLFVGISGAAAAAFOIAKRPENAGKLIIVAFPSFG 302  
Db 304 MDEVVIEISSDEAIFETAKLALKEGLFVGISGAAAAAFOIAKRPENAGKLIIVAFPSFG 363  
QY 303 ERYLSSVLFESVRRRAESM 321  
Db 364 ERYLSSVLFESVRRRAESM 382

## RESULT 8

TS2650  
cysteine synthase (EC 4.2.99.8) precursor, mitochondrion [validated] - Arabidopsis thaliana  
N;Alternate names: O-acetylserine (thiol) lyase  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 24-Oct-2000 #sequence\_revision 24-Oct-2000 #text\_change 03-Nov-2000  
C;Accession: T52650  
R;Jost, R.; Berkowitz, O.; Wirtz, M.; Hopkins, L.; Hawkesford, M.J.; Hell, R.  
Gene 253, 237-247, 2000  
A;Title: Genomic and functional characterization of the oas gene family encoding O-acetylserine synthase  
A;Reference number: Z26157  
A;Accession: T52650  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-387 <JOS>  
A;Cross-references: EMBL:AJ271727, PIDN:CAB71290.1  
A;Experimental source: cultivar Columbia  
C;Genetics:  
A;Gene: oasC  
A;Genome: nuclear  
C;Function:  
A;Description: EC 4.2.99.8 [validated, MUID:20400348]; forms cysteine from O-acetyl-serine  
A;Pathway: cysteine biosynthesis  
C;Superfamily: threonine dehydratase  
C;Keywords: carbon-oxygen lyase; cysteine biosynthesis; mitochondrion  
F;1-36/Domain: transit peptide (mitochondrion) #status predicted <TNP>  
F;37-387/Product: cysteine synthase #status predicted <MAT>

Query Match 78.3%; Score 1271; DB 2; Length 387;  
Best Local Similarity 73.9%; Pred. No. 5e-86;  
Matches 241; Conservative 43; Mismatches 36; Indels 6; Gaps 1;

QY 2 AVERS-----GIADVTTELIGTPTLVYLNKLADGCVARVAAKLELMEPCSSVKRIGYS 55  
Db 58 AVKRETPDGLINADNVSQIGKTPMVLNLSIAKGCVANIAAKLEIMEPCSSVKRIGYS 117  
QY 56 MIADABEKGITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERIILLAF 115  
Db 118 MVTDAEQKGFISPGKSVLVEPTSGNTGIGLAFMAAARGYKLIITMPASMSMERIVLLKAF 177  
QY 116 GAEVLVLTDPKMGKAVOKAEIILAKTPNAYILOQFENPANPKVHYETGPEIWKSDGK 175  
Db 178 GAEVLVLTDPKMGKAVOKAEIILAKTPNAYILOQFENPANPKVHYETGPEIWDPTK 237  
QY 176 IDAFVSGIGTGTGAGKYLEKQNPNIKIGVEPVESPVLSGKPGPHKIQIGAGFIP 235  
Db 238 VDFVAGIGTGTGAGKYLEKQNPNIKIGVEPVESPVLSGKPGPHKIQIGAGFIP 297

QY 236 GVLVNLDEVVOISSDEAIFETAKLALKEGLFVGISGAAAAAFOIAKRPENAGKLI 295  
Db 298 KNLDTKMTDEVVIAISSDEAIFETAKLALKEGLFVGISGAAAAAFOIAKRPENAGKLI 357  
QY 296 AVPSFGERYLSSVLFESVRRRAESM 321  
Db 358 VVPSFGERYLSTPLFQISREVEKM 383

## RESULT 9

T07962  
probable cysteine synthase (EC 4.2.99.8) 1A precursor - Chlamydomonas reinhardtii  
N;Alternate names: O-acetylserine (thiol) lyase; O-acetylserine sulfinhydrilase  
C;Species: Chlamydomonas reinhardtii  
C;Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 21-Jan-2000  
C;Accession: T07962  
R;Ravina, C.G.; Barroso, C.; Vega, J.M.; Gotor, C.  
submitted to the EMBL Data Library, July 1998  
A;Description: Cysteine biosynthesis in Chlamydomonas reinhardtii. Molecular cloning a  
A;Reference number: Z16250  
A;Accession: T07962  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-352 <RAV>  
A;Cross-references: EMBL:AF078693, NID:g3342568, PIDN:AAC27794.1; PID:g3342569  
C;Genetics:  
A;Gene: cys-1A  
C;Superfamily: threonine dehydratase  
C;Keywords: carbon-oxygen lyase; cysteine biosynthesis; phosphoprotein; pyridoxal phos  
F;1-31/Domain: transit peptide (chloroplast) #status predicted <TNP>  
F;32-352/Product: cysteine synthase 1A #status predicted <MAT>  
F;74/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match 78.3%; Score 1270; DB 2; Length 352;  
Best Local Similarity 74.4%; Pred. No. 5.2e-86;  
Matches 241; Conservative 41; Mismatches 42; Indels 0; Gaps 0;

QY 2 AVERSGIADVTTELIGTPTLVYLNKLADGCVARVAAKLELMEPCSSVKRIGYSMIADAE 61  
Db 27 AAVKMIATDVTTELIGTPTMVLNKLAVTGTAKIAAKLEIMEPCSSVKRIGYSMISSAE 86  
QY 62 EKGITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERIILLAFGAEVLV 121  
Db 87 KEGITPGKTVLVEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERIILLAFGAEVLV 146  
QY 122 TDPKMGKAVOKAEIILAKTPNAYILOQFENPANPKVHYETGPEIWKSDGKIDAFVS 181  
Db 147 TDPKMGKAVOKAEIILAKTPNAYILOQFENPANPKVHYETGPEIWSATDGKIDAFVS 206  
QY 182 GGTGTTTGAGKYLEKQNPNIKIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEVN 241  
Db 207 GGTGTTTGAGKYLEKQNPNIKIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEVN 266  
QY 242 LDEVVOISSDEAIFETAKLALKEGLFVGISGAAAAAFOIAKRPENAGKLIIVAFPSF 301  
Db 267 LDEVVOISSDDAIDMARLALIEGLMVGISGAAVQAIAKVPENEGKLVVVVLPF 326  
QY 302 GERYLSSVLFESVRRRAESMTEP 325  
Db 327 GERYLSSVLFQLRDEASKTEP 350

## RESULT 10

S29733  
cysteine synthase (EC 4.2.99.8) B precursor, chloroplast - spinach  
N;Alternate names: O-acetylserine (thiol) lyase; O-acetylserine sulfinhydrilase  
C;Species: Spinacia oleracea (spinach)  
C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 20-Jun-2000  
C;Accession: S29733; S33591; S62218; S23944  
R;Roland, N.; Droux, M.; Lebrun, M.; Douce, R.  
Arch. Biochem. Biophys. 300, 213-222, 1993  
A;Title: O-acetylserine(thiol) lyase from spinach (Spinacia oleracea L.) leaf: cDNA clon  
A;Reference number: S29733; MUID:93143317; PMID:8424655

A:Accession: S29733  
A:Molecule type: mRNA  
A:Residues: 1-383 <ROL>  
A:Cross-references: EMBL:X66860; NID:G312943; PIDN:CAA47329.1; PID:G312944  
A:Experimental source: leaf  
R:Salt, K.; Tatsuguchi, K.; Murakoshi, I.; Hirano, H.  
FEBS Lett. 324, 247-252, 1993  
A:Title: cDNA cloning and expression of cysteine synthase B localized in chloroplasts of  
A:Reference number: S33591; MUID:94009590; PMID:8405359  
A:Accession: S33591  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-11, 'I', 13-333, 'AA', 336-383 <SAL>  
A:Cross-references: GB:D14722; NID:G287441; PIDN:BAA03542.1; PID:G303902  
R:Rolland, N.; Ruffet, M.L.; Job, D.; Douce, R.; Droux, M.  
Eur. J. Biochem. 236, 272-282, 1996  
A:Title: Spinach chloroplast O-acetylserine (thiol)-lyase exhibits two catalytically not  
A:Reference number: S62218; MUID:96184908; PMID:8617276  
A:Accession: S62218  
A:Molecule type: protein  
A:Residues: 53-66; 98-107, 'X', 109-114 <RO2>  
R:Droux, M.; Martin, J.; Sajus, P.; Douce, R.  
Arch. Biochem. Biophys. 295, 379-390, 1992  
A:Title: Purification and characterization of O-acetylserine (thiol) lyase from spinach  
A:Reference number: S23944; MUID:92264737; PMID:1375015  
A:Accession: S23944  
A:Molecule type: protein  
A:Residues: 53-57; 59-60, 'X', 62-65 <DRO>  
A:Experimental source: leaf  
C:Genetics:  
A:Gene: cysk  
A:Genome: nuclear  
C:Function:  
A:Description: catalyzes the sulphydrylation of O-acetylserine in the presence of "sulfi  
A:Pathway: cysteine biosynthesis  
A:Note: last step  
C:Superfamily: threonine dehydratase  
C:Keywords: carbon-oxygen lyase; chloroplast; cysteine biosynthesis; homodimer; phosphor  
P:1-52/Domain: transit peptide (chloroplast) #status predicted <TNP>  
P:53-383/Product: cysteine synthase B #status experimental <MAT>  
P:108/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 77.7%; Score 1261; DB 2; Length 383;  
Best Local Similarity 72.7%; Pred. No. 2.7e-85;  
Matches 232; Conservative 52; Mismatches 35; Indels 0; Gaps 0;

QY 3 VERSGIADVTELLIGTTPVLYLNKLDGCVARVAAKLELMEPCSSVKDRIGYSMTADAE 62  
DB 62 IEGINIAEDVSQIGKTPMVYLNWVSKSVANIAAKLESMEPCSSVKDRIGYSMTADAEQ 121

QY 63 KGLITPGKSVLIBPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILLAFGAELVET 122  
DB 122 KGVITPGKTTLVETSGNTGIGLAFIAARGYKLTITMPASMSMERRVILKAFGAELVLT 181

QY 123 DPAKMGKAVQKABEILAKTPNAVYILQOFENPANPKVHYETTGPEIWKSGDKIDAFVSG 182  
DB 182 DPAKMGKAVQKABEILAKTPNAVYILQOFENPANPKVHYETTGPEIWKSGDKIDAFVAG 241

QY 183 IGTGTTGAGKYLKEQNPNIKLIGVBPVPSVLSGKPGPHKIQIGAGFIPGVLEVN 242  
DB 242 IGTGTTGAGKYLKEQNPNIKLIGVBPVPSVLSGKPGPHKIQIGAGFIPGVLEVN 301

QY 243 IDEVVOISSDEAIEITAKLLAKKEGLFVGISGAAAAAFQIAKRPENAGKLIYAVFPSPG 302  
DB 302 IDEVVOISSDEAIEITAKLLAKKEGLFVGISGAAAAAFQIAKRPENAGKLIYAVFPSPG 361

QY 303 ERYLSSVLFESVRRBAESM 321  
DB 362 ERYLSSVLFESVRRBAESM 380

RESULT 11  
A84870

Cysteine synthase (EC 4.2.99.8) [similarity] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Aug-2002  
C:Accession: A84870  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Tallon, I.  
euss, D.; Nieman, W.C.; White, O.; Eissen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: A84870  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-392 <STO>  
A:Cross-references: GB:AB002093; NID:92281095; PIDN:AAB64031.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g43750  
A:Map position: 2  
C:Superfamily: threonine dehydratase  
C:Keywords: carbon-oxygen lyase

Query Match 77.7%; Score 1261; DB 2; Length 392;  
Best Local Similarity 74.9%; Pred. No. 2.7e-85;  
Matches 239; Conservative 41; Mismatches 39; Indels 0; Gaps 0;

QY 3 VERSGIADVTELLIGTTPVLYLNKLDGCVARVAAKLELMEPCSSVKDRIGYSMTADAE 62  
DB 70 VEGINIAEDVSQIGKTPMVYLNWVSKSVANIAAKLESMEPCSSVKDRIGYSMTADAE 129

QY 63 KGLITPGKSVLIBPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILLAFGAELVET 122  
DB 130 KGLITPGKSVLIBPTSGNTGIGLAFIAARGYKLTITMPASMSMERRVILKAFGAELVLT 189

QY 123 DPAKMGKAVQKABEILAKTPNAVYILQOFENPANPKVHYETTGPEIWKSGDKIDAFVSG 182  
DB 190 DPAKMGKAVQKABEILAKTPNAVYILQOFENPANPKVHYETTGPEIWKSGDKIDAFVAG 249

QY 183 IGTGTTGAGKYLKEQNPNIKLIGVBPVPSVLSGKPGPHKIQIGAGFIPGVLEVN 242  
DB 250 IGTGTTGAGKYLKEQNPNIKLIGVBPVPSVLSGKPGPHKIQIGAGFIPGVLEVN 309

QY 243 IDEVVOISSDEAIEITAKLLAKKEGLFVGISGAAAAAFQIAKRPENAGKLIYAVFPSPG 302  
DB 310 IDEVVOISSDEAIEITAKLLAKKEGLFVGISGAAAAAFQIAKRPENAGKLIYAVFPSPG 369

QY 303 ERYLSSVLFESVRRBAESM 321  
DB 370 ERYLSSVLFESVRRBAESM 388

RESULT 12  
T47800  
cysteine synthase (EC 4.2.99.8) F24G16.30 [similarity] - Arabidopsis thaliana  
N:Alternate names: protein F24G16.30  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 03-Nov-2000  
C:Accession: T47800  
R:D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.  
submitted to the Protein Sequence Database, February 2000  
A:Reference number: Z24477  
A:Accession: T47800  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-442 <DAN>  
A:Cross-references: EMBL:AL138647  
A:Experimental source: cultivar Columbia; BAC clone F24G16  
C:Genetics:  
A:Map position: 3  
A:Introns: 119/3; 156/2; 175/3; 263/3; 309/3; 327/1; 353/3; 373/3; 400/3  
A:Note: F24G16.30  
C:Superfamily: threonine dehydratase  
C:Keywords: carbon-oxygen lyase; phosphoprotein; pyridoxal phosphate  
P:154/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 77.3%; Score 1255; DB 2; Length 442;  
Best Local Similarity 71.3%; Pred. No. 8.9e-85;  
Matches 241; Conservative 43; Mismatches 36; Indels 18; Gaps 2;

QY 2 AVER8-----GIADVTLEIGKTPLVYLNKLDGCVARVAAKLEIMPCSSVKDRIGYS 55  
DB 101 AVKRETPDGLNIADNVSQLIKTPMVTYLNIAKGCYANIAAKLEIMEPCSSVKDRIGYS 160  
QY 56 MIADAEKGLITPKSVLIEPTSGNTGIGLAFMAARGYKLIITMPASMSLERIILIAF 115  
DB 161 MWTDAEQKGFISPGKSVLIEPTSGNTGIGLAFIAASRGYKLIITMPASMSMERVLIAF 220  
QY 116 GAEVLVTPAKMGKGAVOAKAEIILAKTPNAVILQGFENPANPKVHYETTGPEIWKSDGK 175  
DB 221 GAEVLVTPAKMGKGAVOAKAEIILAKTPNAVILQGFENPANPKVHYETTGPEIWKSDGK 280  
QY 176 IDAFVSGIGTGTITGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIP 235  
DB 281 VDFVAGIGTGTITGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIP 340  
QY 236 GVLVNLDEVOISSDEALETAKLALKEGLFVGISSGAAAAAFQIAKRPENAGKLIV 295  
DB 341 KNLQKIMEVIAISSEALIBAKQALKEGLFVGISSGAAAAAFQIAKRPENAGKLIV 400  
QY 296 -----AVFPFGRYLLSVLFESVRRBAESM 321  
DB 401 VTNESMNAVFQVFPFGRYLLSVLFESVRRBAESM 438

## RESULT 13

S48694  
cysteine synthase (EC 4.2.99.8) isoform 5-8, cytosolic - Arabidopsis thaliana  
N:Alternate names: O-acetylserine (thiol) lyase  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 16-Feb-1995 #sequence\_revision 19-Jan-1996 #text\_change 22-Jun-1999  
C:Accession: S48694  
R:Hell, R.; Bork, C.; Bogdanova, N.; Frolov, I.; Hauschild, R.  
FEBS Lett. 351, 257-262, 1994

A:Title: Isolation and characterization of two cDNAs encoding for compartment specific  
A:Reference number: S48694; MUID:94364481; PMID:8082776  
A:Accession: S48694  
A:Molecule type: mRNA  
A:Residues: 1-315 <HEU>  
A:Cross-references: EMBL:X80376; MID:9602329; PIDN:CAA56593.1; PID:9560130  
C:Superfamily: threonine dehydratase  
C:Keywords: carbon-oxygen lyase; cysteine biosynthesis; cytosol; homodimer; phosphoprote  
F:46/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 77.1%; Score 1252; DB 2; Length 315;  
Best Local Similarity 80.3%; Pred. No. 9.5e-85;  
Matches 245; Conservative 28; Mismatches 30; Indels 2; Gaps 1;

QY 6 SGIADVTLEIGKTPLVYLNKLDGCVARVAAKLEIMPCSSVKDRIGYSMTADAEKGL 65  
DB 3 SRIADVTLEIGKTPLVYLNKLDGCVARVAAKLEIMPCSSVKDRIGYSMTADAEKGL 62  
QY 66 ITPGKSVLIEPTSGNTGIGLAFMAARGYKLIITMPASMSLERIILIAFCAELVLTDP 125  
DB 63 IKPESVLIPTSGNTGIGLAFMAARGYKLIITMPASMSLERIILIAFCAELVLTDP 122  
QY 126 KGMKGAVOAKAEIILAKTPNAVILQGFENPANPKVHYETTGPEIWKSDGKIDAFVSGIGT 185  
DB 123 KGMKGAVOAKAEIILAKTPNAVILQGFENPANPKVHYETTGPEIWKSDGKIDAFVSGIGT 182  
QY 186 GGTITGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLVNLDE 245  
DB 183 G--ITGAGSILKNRTANVLYGVEPVESAILIGKPGPHKIQIGAGFIPVNLVNLDE 240  
QY 246 VVQISSDEALETAKLALKEGLFVGISSGAAAAAFQIAKRPENAGKLIVAVFPFSGERY 305  
DB 241 VVQISSDEALETAKLALKEGLFVGISSGAAAAAFQIAKRPENAGKLIVAVFPFSGERY 300

QY 306 LSSVL 310  
DB 301 LSTVL 305

## RESULT 14

A43407  
cysteine synthase (EC 4.2.99.8) precursor - pepper  
N:Alternate names: O-acetylserine (thiol) lyase; O-acetylserine sulfinhydratase  
C:Species: Capsicum annuum (pepper)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A43407; S24637  
R:Romer, S.; d'Harlingue, A.; Camara, B.; Schantz, R.; Kuntz, M.  
J. Biol. Chem. 267, 17966-17970, 1992  
A:Title: Cysteine synthase from Capsicum annuum chromatoplasts. Characterization and cDN  
A:Reference number: A43407; MUID:92388158; PMID:1381358  
A:Accession: A43407  
A:Molecule type: mRNA  
A:Residues: 1-374 <ROM>  
A:Cross-references: GB:X64874; MID:917943; PIDN:CAA46086.1; PID:917944  
A>Note: sequence extracted from NCBI backbone (NCBI:112876)  
C:Genetics:  
A:Genome: nuclear  
C:Superfamily: threonine dehydratase  
C:Keywords: carbon-oxygen lyase; chloroplast; cysteine biosynthesis; phosphoprotein; p  
F:108/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 77.0%; Score 1249; DB 1; Length 374;  
Best Local Similarity 75.2%; Pred. No. 2e-84;  
Matches 240; Conservative 38; Mismatches 31; Indels 10; Gaps 1;

QY 3 VERSIADVTLEIGKTPLVYLNKLDGCVARVAAKLEIMPCSSVKDRIGYSMTADAE 62  
DB 62 IEGNIAEDVTQIIGNTPMVTYLNIAKGCYANIAAKLEIMEPCSSVKDRIGYSMTADAE 121  
QY 63 KGLITPKSVLIEPTSGNTGIGLAFMAARGYKLIITMPASMSLERIILIAFCAELVLT 122  
DB 122 KGLISPGKTVLIEPTSGNTGIGLAFIAASRGYKLIITMPASMSLERIILIAFCAELVLT 181  
QY 123 DPAKMGKGAVOAKAEIILAKTPNAVILQGFENPANPKVHYETTGPEIWKSDGKIDAFVSG 182  
DB 182 DPAKMGKGAVOAKAEIILAKTPNAVILQGFENPANPKVHYETTGPEIWKSDGKIDAFVSG 241  
QY 183 IGTGTTGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLVNL 242  
DB 242 IGTGTTGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLVNL 291  
QY 243 LDRVQISSDEALETAKLALKEGLFVGISSGAAAAAFQIAKRPENAGKLIVAVFPFSG 302  
DB 292 LDRVQISSDEALETAKLALKEGLFVGISSGAAAAAFQIAKRPENAGKLIVAVFPFSG 351  
QY 303 BRYLSSVLPESVRRBAESM 321  
DB 352 BRYLSSVLPESVRRBAESM 370

## RESULT 15

S49586  
cysteine synthase (EC 4.2.99.8) ACS1 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 27-Jan-1995 #sequence\_revision 24-Feb-1995 #text\_change 22-Jun-1999  
C:Accession: S49586  
R:Hesse, H.; Altmann, T.  
submitted to the EMBL Data Library, September 1994  
A:Description: Molecular cloning of a cysteine synthase cDNA from Arabidopsis thaliana  
A:Reference number: S49586  
A:Accession: S49586  
A:Molecule type: mRNA  
A:Residues: 1-324 <HBS>  
A:Cross-references: EMBL:X81697; MID:9572518; PIDN:CAA57343.1; PID:9572519  
C:Genetics:  
A:Gene: ACS1  
C:Superfamily: threonine dehydratase



C:Keywords: carbon-oxygen lyase; cysteine biosynthesis; phosphoprotein; pyridoxal phosph  
F:49/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match 75.7%; Score 1229; DB 2; Length 324;

Best Local Similarity 75.9%; Pred. No. 4.8e-83;

Matches 246; Conservative 34; Mismatches 38; Indels 6; Gaps 4;

QY	6	SGIAKDVTELGKTPVLYLNKLA-DGCVARVAAKLEMEPCS--SYKDRIGYMTADAER	62
Db	3	SRIAKDVTELGKTPVLYLNKLA-DGCVARVAAKLEMEPCS--SYKDRIGYMTADAER	62
QY	63	KGLITPGKSVLIEPTSGNTG-IGLAFMAARGYKLTITMPASMSLERIILLAFGAEVL	121
Db	63	KGLITPGKSVLIEPTSGNTG-IGLAFMAARGYKLTITMPASMSLERIILLAFGAEVL	122
QY	122	TDPKGNKGAQVQAEIILAKTNPAYILQGFENPANPKVHETTGPEIWKSGDGIIDAFVS	181
Db	123	TDPKGNKGAQVQAEIILAKTNPAYILQGFENPANPKVHETTGPEIWKSGDGIIDAFVS	182
QY	182	GIGTGGTITGAGKYLKEQNPNIKLIGVEPVESPVLSGGRPGPHKIQIGAGFIPGVLEVN	241
Db	183	GIGTGGTITGAGKYLKEQNPNIKLIGVEPVESPVLSGGRPGPHKIQIGAGFIPGVLEVN	242
QY	242	LEDEVVQISSDEAIEETAKLALKEGLFVGISSGAAAAAFQIAKRPENAGKLIYAVFPSP	301
Db	243	LEDEVVQISSDEAIEETAKLALKEGLFVGISSGAAAAAFQIAKRPENAGKLIYAVFPSP	300
QY	302	GERTLSSVLFESVRRRABSMTEP	325
Db	301	GERTLSSVLFESVRRRABSMTEP	324

Search completed: May 18, 2004, 13:09:04  
Job time : 21 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2004, 13:02:39 ; Search time 17 Seconds  
(without alignments)  
995.459 Million cell updates/sec

Title: US-09-931-457A-31  
Perfect score: 1623  
Sequence: 1 MAVERSGIAKDVTETLICKTP.....LSSVLFESVRRKAESEMTFEP 325

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	1453	89.5	325 1	CYSK_CITLA Q43317 citrullus 1
2	1435	88.4	325 1	CYSK_SOLTU O81154 solanum tub
3	1431	88.2	325 1	CYSK_SPIOL Q00834 spinacia ol
4	1417	87.3	322 1	CYK1_BRAJU O23733 brassica ju
5	1409	86.8	324 1	CYK2_BRAJU O23735 brassica ju
6	1384	85.3	322 1	CYSK_ARATH P47998 arabidopsis
7	1378	84.9	325 1	CYSK_MAIZE P80608 zea mays (m
8	1377	84.8	325 1	CYK2_ORYSA O9x6a8 cryza sativ
9	1364	84.0	325 1	CYSK_WHEAT P38076 triticum ae
10	1356	83.5	321 1	CYK1_ORYSA O9x6a6 cryza sativ
11	1314	81.0	386 1	CYSL_SOLTU O81155 solanum tub
12	1271	78.3	383 1	CYSL_SPIOL P32260 spinacia ol
13	1271	78.3	424 1	CYSM_ARATH Q43725 arabidopsis
14	1261	77.7	392 1	P47999 arabidopsis
15	1249	77.0	374 1	CYSL_CAPAN P31300 capsicum an
16	1045	64.4	404 1	CYSN_ARATH O22682 arabidopsis
17	971	59.8	312 1	CYSK_SYNY3 P73410 synechocyst
18	949.5	58.5	310 1	CYSK_MYCTU P95230 mycobacteri
19	940.5	57.9	310 1	CYSK_MYCLE O32978 mycobacteri
20	869.5	53.6	329 1	SRPG_SYNP7 Q59966 synechococ
21	865.5	53.3	322 1	CYSK_SALTY P12674 salmoneilla
22	859	52.9	307 1	CYSK_BACSU P37877 bacillus su
23	857.5	52.8	322 1	CYSK_BACOL P11096 escherichia
24	816.5	50.3	316 1	CYSK_HAEIN P45040 haemophilus
25	815	50.2	310 1	CYSK_STAM Q99w90 staphylococ
26	801	49.4	310 1	CYSK_STAP Q8cmc6 staphylococ
27	769.5	47.4	322 1	CYSK_BUCAP Q8ka48 buchnera ap
28	765.5	47.2	315 1	CYSK_BUCAI P57171 buchnera ap
29	737.5	45.4	311 1	CYSM_BACSU Q59447 bacillus su
30	701	43.2	307 1	CYSK_FLASP Q59447 flavobacter
31	682	42.0	299 1	CYSM_CAMJCB P71128 campylobact
32	611.5	37.7	303 1	CYSM_SALTY P29848 salmonella
33	610.5	37.6	303 1	CYSM_ECOLI P16703 escherichia

34	598.5	36.9	327 1	CYSM_AQUAE O67507 aquifex aeo
35	585.5	36.1	306 1	CYSM_HELUP P56067 helicobacte
36	577.5	35.6	305 1	CYSM_HELUP Q9znmw6 helicobacte
37	550	33.9	550 1	CBS_HUMAN P35520 homo sapien
38	545	33.6	560 1	CBS_RAT P32232 rattus norv
39	544.5	33.5	497 1	CBS_DICTDI P46794 dictyosteli
40	541.5	33.4	507 1	CBS_YEAST P32582 saccharomyc
41	483	29.8	323 1	CYSM_MYCTU Q10624 mycobacteri
42	463	28.5	393 1	CYSK_YEAST P53206 saccharomyc
43	459	28.3	371 1	CYSK_EMENT P50867 emericaella
44	433	26.7	155 1	CYSK_STAHA Q59918 staphylococ
45	432	26.6	395 1	CYSK_SCHPO P87131 schizosacch

## ALIGNMENTS

RESULT 1  
CYSK\_CITLA STANDARD; PRT; 325 AA.  
ID CYSK\_CITLA  
AC Q43317;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cysteine synthase (EC 2.5.1.47) (Beta-pyrazolylalanine synthase)  
DE (Beta-PA/CSase) (EC 2.5.1.51) (L-mimosine synthase) (EC 2.5.1.52) (O-acetylserine sulphydrylase) (O-acetylserine (Thiol)-lyase) (CSase) (OAS-TL).  
OS Citrullus lanatus (Watermelon) (Citrullus vulgaris).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Cucurbitales; Cucurbitaceae; Citrullus.  
OX NCBI\_TaxID=3654;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Seedling;  
RX MEDLINE=94316193; PubMed=8041362;  
RA Noji M., Murakoshi I., Saito K.;  
RT "Molecular cloning of a cysteine synthase cDNA from Citrullus  
RT vulgaris (watermelon) by genetic complementation in an Escherichia  
RT coli Cys-auxotroph.";  
RT Mol. Gen. Genet. 244:57-66 (1994).  
RL [2]  
RN [2]  
RP FUNCTION.  
RX MEDLINE=94107305; PubMed=8280125;  
RA Noji M., Murakoshi I., Saito K.;  
RT "Evidence for identity of beta-pyrazolealanine synthase with cysteine  
RT synthase in watermelon: formation of beta-pyrazole-alanine by cloned  
RT cysteine synthase in vitro and in vivo.";  
RT Biochem. Biophys. Res. Commun. 197:1111-1117 (1993).  
RL [3]  
RN [3]  
RP FUNCTION.  
RX MEDLINE=97165913; PubMed=9013806;  
RA Saito K., Kimura N., Ikegami F., Noji M.;  
RT "Production of plant non-protein amino acids by recombinant enzymes of  
RT sequential biosynthetic reactions in bacteria.";  
RT Biol. Pharm. Bull. 20:47-53 (1997).  
RL [1-]  
CC -1- FUNCTION: Produces L-cysteine from O-acetyl-L-serine and hydrogen  
CC sulfide. Can also use pyrazole and 3,4-dihydroxypyridine instead  
CC of the hydrogen sulfide to produce two plant specific non-  
CC protein amino acids beta-pyrazolylalanine and L-mimosine.  
CC -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +  
CC acetate.  
CC -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + pyrazole = 3-(pyrazol-  
CC 1-yl)-L-alanine + acetate.  
CC -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + 3,4-dihydroxypyridine =  
CC 3-(3,4-dihydroxypyridin-1-yl)-L-alanine + acetate.  
CC -1- COFACTOR: Pyridoxal phosphate.  
CC -1- PATHWAY: Cysteine biosynthesis.  
CC -1- SUBUNIT: Homodimer (By similarity).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-  
CC synthase family.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

DR   EMBL; D28777; BAA05965.1; -.
DR   PIR; S46438; S46438.
DR   HSSP; P12674; 1OAS.
DR   InterPro; IPR001926; B6_enzyme_beta.
DR   InterPro; IPR001216; Cys_synthase_BS.
DR   InterPro; IPR005859; Cys_synthK.
DR   InterPro; IPR005856; Cys_synthKM.
DR   Pfam; PF00291; PALP; 1.
DR   TIGRFAMs; TIGR01139; cysK; 1.
DR   TIGRFAMs; TIGR01136; cysKM; 1.
DR   PROSITE; PS00901; CYS_SYNTHASE; 1.
DR   Transferrase, Cysteine biosynthesis; Pyridoxal phosphate.
KM   BINDING 49 49 PYRIDOXAL PHOSPHATE (BY SIMILARITY)
FT   DOMAIN 275 280 POLY-ALA.
SQ   SEQUENCE 325 AA; 34342 MW; 4A91E6F8AFB5F539 CRC64;

```

Query Match	89.5%	Score 1453;	DB 1;	Length 325;
Best Local Similarity	87.1%;	Pred. No. 2.1e-92;		
Matches 263;	Conservative	24;	Mismatches 18;	Indels 0;
				Gaps 0;

QY	1	MAVERSGIAKDVTBELIGKTPVLYTNKLADGCVARVAAKLEIMEPCSSVKORIGYSMTADA	60
Db	1	MADAKSTIAKDVTBELIGNTPLVYTLNRVVDGCVARVAAKLEIMEPCSSVKDRIGYSMTSBA	60
QY	61	BEKGLITPKGSVLIPTSGNTGIGLAFMAARGYKLIITMPASMSLERRIILAFGAEV	120
Db	61	ENKGLITPGESVLIPTSGNTGIGLAFIAAARGYRIIICMPASMSLERRITILRAFGEV	120
QY	121	LTPDAKMGKAVOKABEILAKTPNAYILOQFENPANPKVHYETTGTPEIWRGSDGKIDAV	180
Db	121	LTPDARGMKGAVOKABEIKAKTPNSYILOQFENPANPKIHETTGTPEIWRGSGGKIDALV	180
QY	181	SGIGTGGITGAGKYLKEONPNIKLIGVEPVEPSVLSGGKPGPHKIQIGIGAGFI PGVLAV	240
Db	181	SGIGTGGITGAGKYLKEONPNIKLYGVEPVEPSAIIISGGKPGPHKIQIGIGAGFI PGVLAV	240
QY	241	NILDEVQISSDEALISTAKLALKEGLFVGISSGAAAAAAAFQIAKRPENAGKLIIVAVPFS	300
Db	241	NILDEVQIVSSSEESIETAKLALKEGLLVGISSGAAAAAAARIKRPENAGKLIIVAVPFS	300
QY	301	FGERYLSSVLFESVRRABSMTFEP	325
Db	301	FGERYLSTVLFESVKRETENMFEP	325

RESULT 2

ID	CYSK SOLTU	STANDARD;	PRT;	325 AA.
----	------------	-----------	------	---------

DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cysteine synthase (BC 2.5.1.47) (O-acetylserine sulphydrylase)  
DE (O-acetylserine (Thiol)-lyase) (CSase A) (CS-A) (OAS-TL A) .  
OS Solanum tuberosum (Potato) .  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Lamiales; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4113;

RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Berolina;  
RA Hesse H., Hoefgen R.;  
RT # Isolation of cDNAs encoding cytosolic and plastidic cysteine synthase

RT isoforms from *Solanum tuberosum*.";  
 RL (in) Plant Gene Register PGR98-057.  
 CC -| CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +  
 CC acetate.  
 CC -| COFACTOR: Pyridoxal phosphate.  
 CC -| PATHWAY: Cysteine biosynthesis.  
 CC -| SUBUNIT: Homodimer (By similarity).  
 CC -| SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -| SIMILARITY: Belongs to the cysteine synthase/cystathionine beta  
 CC synthase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR PIR; T07001; T07001.  
DR HSEB; PI2674; 1FCJ.  
DR InterPro; IPR001926; B6\_enzyme\_beta.  
DR InterPro; IPR001216; Cys\_synthase\_BS.  
DR InterPro; IPR005859; Cys\_synthK.  
DR InterPro; IPR005856; Cys\_synthKX.  
DR Pfam; PF00291; PALP; 1.  
DR TIGRFAMS; TIGR01139; cysK; 1.  
DR TIGRFAMS; TIGR01136; cysKM; 1.  
DR PROSITE; PS00901; CYS\_SYNTHASE; 1.  
KW Transferase; Cysteine\_biosynthesis; Pyridoxal phosphate.  
FT BINDING 49 49 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
FT DOMAIN 275 280 POLY-ALA.  
SQ SEQUENCE 325 AA; 34341 MW; 86119D0466C647BC CRC64;

Query Match	88.4%;	Score 1435;	DB 1;	Length 325;
Best Local Similarity	85.2%;	Pred. No. 3.5e-91;		
Matches 277;	Conservative 26;	Mismatches 22;	Indels 0;	Gaps 0;

Qy	1	MAVERSGIAKDVTTELIGKTPLVYLNKLDAGCVARVAAKLELMEPCSSVKDRIGYSMIADA	60
Db	1	MAGEKIGIAKDVTTELIGNTPLVYINNVDGCVARVAAKLESMEPCCSSVKDRIGYSMITDA	60
Qy	61	BEKGILTPGKSVLIEPTSGNTGIGIAFMAAARGYKLIITMPASMSLEBRILLAFGAEV	120
Db	61	BEKGLTPGESVLIERTSGNTGVGIAFMAAARGYKLIITMPSSMSLEBRILLRGFRSEV	120
Qy	121	LTPDPAKGKGAVOQKABRIIAKTPNAYILQOFENPANPKVHYETTGPEIMKSGDGKIDAFV	180
Db	121	LTPDPAKGKGAISKABEIKAKTPNBYILQOFENPANPKIHYETTGPEIMKSGNGKVDALA	180
Qy	181	SGIGTGTITGAGKYLKEQNPNIKLIGYEPVESPVLSGKPGPHKIQGIGAGFPGVLEV	240
Db	181	SGIGTGTITGSGKYLREQNPNVKLYGEPVESPAIILSGKPGPHKIQGIGAGFPGVLEV	240
Qy	241	NLDEWQIISDEAIETAKLLAKKEGLEVGIISSGAAAAAAFOIAKRPENAGKLIIVAFPS	300
Db	241	NLIDWQVQSSSESIEMAKLLAKKEGLLYGIISSGAAAAAAIKVAKRPENAGKLIIVIFPS	300
Qy	301	FGERYLISSVLPESVRRBAESMTPEP	325
Db	301	FGERYLISSVLPETVRRBAENMTVEP	325

### RESULT 3

ID	CYCK SPIOL	STANDARD;	PRT;	325 AA.
1	1	1	1	1
2	2	2	2	2
3	3	3	3	3
4	4	4	4	4
5	5	5	5	5
6	6	6	6	6
7	7	7	7	7
8	8	8	8	8
9	9	9	9	9
10	10	10	10	10
11	11	11	11	11
12	12	12	12	12
13	13	13	13	13
14	14	14	14	14
15	15	15	15	15
16	16	16	16	16
17	17	17	17	17
18	18	18	18	18
19	19	19	19	19
20	20	20	20	20
21	21	21	21	21
22	22	22	22	22
23	23	23	23	23
24	24	24	24	24
25	25	25	25	25
26	26	26	26	26
27	27	27	27	27
28	28	28	28	28
29	29	29	29	29
30	30	30	30	30
31	31	31	31	31
32	32	32	32	32
33	33	33	33	33
34	34	34	34	34
35	35	35	35	35
36	36	36	36	36
37	37	37	37	37
38	38	38	38	38
39	39	39	39	39
40	40	40	40	40
41	41	41	41	41
42	42	42	42	42
43	43	43	43	43
44	44	44	44	44
45	45	45	45	45
46	46	46	46	46
47	47	47	47	47
48	48	48	48	48
49	49	49	49	49
50	50	50	50	50
51	51	51	51	51
52	52	52	52	52
53	53	53	53	53
54	54	54	54	54
55	55	55	55	55
56	56	56	56	56
57	57	57	57	57
58	58	58	58	58
59	59	59	59	59
60	60	60	60	60
61	61	61	61	61
62	62	62	62	62
63	63	63	63	63
64	64	64	64	64
65	65	65	65	65
66	66	66	66	66
67	67	67	67	67
68	68	68	68	68
69	69	69	69	69
70	70	70	70	70
71	71	71	71	71
72	72	72	72	72
73	73	73	73	73
74	74	74	74	74
75	75	75	75	75
76	76	76	76	76
77	77	77	77	77
78	78	78	78	78
79	79	79	79	79
80	80	80	80	80
81	81	81	81	81
82	82	82	82	82
83	83	83	83	83
84	84	84	84	84
85	85	85	85	85
86	86	86	86	86
87	87	87	87	

DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cysteine synthase (EC 2.5.1.47) (O-acetylserine sulfinhydrilase)  
DE (O-acetylserine (Thiol)-lyase) (CSase A) (OAS-TL A).



OS Spinacia oleracea (Spinach).  
OC Burkaria; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllales; Amarantaceae; Spinacia.  
ON NCBI\_TaxID=3562;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 5-11; 71-90; 163-182 AND 246-262.  
RC STRAIN=cv. Parade; TISSUE=Leaf;  
RX MEDLINE=93390392; PubMed=1518833;  
RA Saito K., Miura N., Yamazaki M., Hirano H., Murakoshi I.;  
RT "Molecular cloning and bacterial expression of cDNA encoding a plant  
cysteine synthase.";  
RL Proc. Natl. Acad. Sci. U.S.A. 89:8078-8082(1992).  
RN [2]  
RP MUTAGENESIS OF LYSINE RESIDUES.  
RX MEDLINE=93345669; PubMed=8344414;  
RA Saito K., Kurosawa M., Murakoshi I.;  
RT "Determination of a functional lysine residue of a plant cysteine  
synthase by site-directed mutagenesis, and the molecular evolutionary  
implications.";  
RL FEBS Lett. 328:111-114(1993).  
CC -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +  
acetate.  
CC -1- COFACTOR: Pyridoxal phosphate.  
CC -1- PATHWAY: Cysteine biosynthesis.  
CC -1- SUBUNIT: Homodimer.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- TISSUE SPECIFICITY: Leaves and roots.  
CC -1- PTM: The N-terminus is blocked.  
CC -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-  
synthase family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; D10476; BAA01279.1; -  
DR PIR; S35094; S35094.  
DR HSSP; P12674; IOAS.  
DR InterPro; IPR001926; B6 enzyme beta.  
DR InterPro; IPR001216; Cys\_synthase\_BS.  
DR InterPro; IPR005859; Cys\_synthK.  
DR InterPro; IPR005856; Cys\_synthKM.  
DR Pfam; PF00291; PALP; 1.  
DR TIGRFAMS; TIGR01139; cysK; 1.  
DR TIGRFAMS; TIGR01136; cysKM; 1.  
DR PROSITE; PS00901; CYS\_SYNTHASE; 1.  
DR Transferase; Cysteine biosynthesis; Pyridoxal phosphate.  
FT BINDING 49 49 PYRIDOXAL PHOSPHATE.  
FT DOMAIN 275 280 POLY-ALA.  
SQ SEQUENCE 325 AA; 34186 MW; 161B46F7B670DEB6 CRC64;  
QY 1 MAVERSGIAKDVTELGKTPVLYLNKLADGCVARVAAKLELMEPCSSVKDRIGYSMLADA 60  
DB 1 MVEEKAFAIAKDVTELGKTPVLYLNKLADGCVARVAAKLELMEPCSSVKDRIGYSMLTDA 60  
QY 61 EEKGIIITPGKSVLIEPTSGNTGIGLAFAAARGYKLIITPASMSLERRIITLAFGALV 120  
DB 61 EKSGLITPGESVLIPTSGNTGIGLAFAAARGYKLIITPASMSLERRITLAFGALV 120  
QY 121 LTDPKAGKAVOKAEILAKTPNAYILQQFENPANPKVHYETTGPETWKSDDKIDAFV 180  
DB 121 LTDPKAGKAVOKAEILAKTPNAYILQQFENPANPKVHYETTGPETWKSDDKIDAFV 180  
QY 181 SGIGTGTTGAGKYLKEQNPNIKLIGVEPVESVLSGKGPHKIQIGAGFIQVLEV 240

DB 181 SGIGTGTTGAGKYLKEQNPNIKLIGVEPVESVLSGKGPHKIQIGAGFIQVLDV 240  
QY 241 NLDEVQISSDEALETAKLALKEGLFVGISSGAAAAAFQIAKRPENAGKLIIVAFPS 300  
DB 241 NIIDEVQISSDESIEMAKLALKEGLVGISSGAAAAAIAKVAKRPENAGKLIIVAFPS 300  
QY 301 FGERYLSVLFESVVRBAESMTPE 324  
DB 301 FGERYLSVLFESVVRBAESMTPE 324  
RESULT 4  
CYKL BRAUT STANDARD; PRT; 322 AA.  
ID CYKL BRAUT  
AC O23733;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cysteine synthase (EC 2.5.1.47) (O-acetylserine sulphydrylase)  
DE (O-acetylserine (Thiol)-lyase) (CSase) (OAS-TL) (OAS-TL4).  
OS Brassica juncea (leaf mustard) (Indian mustard).  
OC Burkaria; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC Euphorbia; Brassicales; Brassicaceae; Brassica.  
ON NCBI\_TaxID=3707;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Vitasso; TISSUE=Root;  
RX MEDLINE=98281577; PubMed=9620267;  
RA Schaefer H.J., Haag-Korner A., Rausch T.;  
RT "cDNA cloning and expression analysis of genes encoding GSH synthesis  
in roots of the heavy-metal accumulator Brassica juncea L.: evidence  
for Cd-induction of a putative mitochondrial gamma-glutamylcysteine  
synthetase isoform.";  
RL Plant Mol. Biol. 37:87-97(1998).  
CC -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +  
acetate.  
CC -1- COFACTOR: Pyridoxal phosphate (By similarity).  
CC -1- PATHWAY: Cysteine biosynthesis.  
CC -1- SUBUNIT: Homodimer (By similarity).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-  
synthase family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; Y10845; CAA71798.1; -  
DR HSSP; P12674; IOAS.  
DR InterPro; IPR001926; B6 enzyme beta.  
DR InterPro; IPR001216; Cys\_synthase\_BS.  
DR InterPro; IPR005859; Cys\_synthK.  
DR InterPro; IPR005856; Cys\_synthKM.  
DR Pfam; PF00291; PALP; 1.  
DR TIGRFAMS; TIGR01139; cysK; 1.  
DR TIGRFAMS; TIGR01136; cysKM; 1.  
DR PROSITE; PS00901; CYS\_SYNTHASE; 1.  
DR Transferase; Cysteine biosynthesis; Pyridoxal phosphate.  
FT BINDING 46 46 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
FT DOMAIN 272 277 POLY-ALA.  
SQ SEQUENCE 322 AA; 33902 MW; 28DA67631EF0ABF2 CRC64;  
QY 6 SGIAKDVTELGKTPVLYLNKLADGCVARVAAKLELMEPCSSVKDRIGYSMLADAEEKGL 65  
Query Match 87.3%; Score 1417; DB 1; Length 322;  
Best Local Similarity 85.0%; Pred. No. 6e-90;  
Matches 271; Conservative 30; Mismatches 18; Indels 0; Gaps 0;

```

DB 3 SRIADVTLLIGNTPLVYLNVAEGCVGRVAAKLEMEPCSSVKDRIGFSMISDAOKGL 62
QY 66 ITPGKSVLIBPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILLAFGAEVLVLTDP 125
DB 63 IKPGESVLIBPTSGNTGIGLAFAAARGYKLIITMPASMSVERRIILLAFGAEVLVLTDP 122
QY 126 KGMKGAIVAKAEIILAKTPNAYILQOFENPANPKHYETTGPETWKSDDKIDAFVSGIGT 185
DB 123 KGMKGAIVAKAEIILAKTPNAYILQOFENPANPKHYETTGPETWKSDDKIDAFVSGIGT 182
QY 186 GGTITGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEVNLLD 245
DB 183 GGTITGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEVNLLD 242
QY 246 VQOISSDEAIEITAKLALKEGLFVGISGAAAAAFOIAKRPENAGKLIYAVPSPFGERY 305
DB 243 VQOISSDESIDMARLLAKREGLLVGISSGAAAAAIIKAKRPENAGKLIYAVPSPFGERY 302
QY 306 LSSVLFESVRRRAESMTFE 324
DB 303 LSTVLFDAITRKBAESMTFQ 321

RESULT 5
CYK2_BRAUT STANDARD; PRT; 324 AA.
ID CYK2_BRAUT STANDARD; PRT; 324 AA.
AC 023735;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cysteine synthase (EC 2.5.1.47) (O-acetylserine sulphydrylase)
DE (O-acetylserine (thiol)-lyase) (CSase) (OAS-TL) (OAS-TL6).
OS Brassica juncea (leaf mustard) (Indian mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_Taxid=3707;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Vittasso; TISSUE=Root;
RX MEDLINE=98281577; PubMed=9620267;
RA Schaefer H.J., Haag-Korner A., Rausch T.;
RT "cDNA cloning and expression analysis of genes encoding GSH synthesis
RT in roots of the heavy-metal accumulator Brassica juncea L.: evidence
RT for Cd-induction of a putative mitochondrial gamma-glutamylcysteine
RT synthetase isoform."
RT Plant Mol. Biol. 37:87-97 (1998).
CC -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +
CC acetate.
CC -1- COFACTOR: Pyridoxal phosphate (By similarity).
CC -1- PATHWAY: Cysteine biosynthesis.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-
CC synthase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y10847; CAA71800.1; -.
DR HSSP: P12674; 1OAS.
DR InterPro: IPR001926; B6_enzyme_beta.
DR InterPro: IPR001216; Cys_synthase_BS.
DR InterPro: IPR005859; Cys_synthase.
DR InterPro: IPR005856; Cys_synthase.
DR Pfam: PF00291; PALP; 1.
DR TIGRFAMs: TIGR01139; cybK; 1.

```

```

DR TIGRFAMs: TIGR01136; cybK; 1.
DR PROSITE: PS00901; CYS_SYNTHASE; 1.
KM Transferrase; Cysteine biosynthesis; Pyridoxal phosphate.
FT BINDING 48 48 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT DOMAIN 274 279 POLY-ALA.
SQ SEQUENCE 324 AA; 34132 MW; 091B5B2C0BED1EAI CRC64;

Query Match 86.8%; Score 1409; DB 1; Length 324;
Best Local Similarity 84.1%; Pred. No. 2.1e-89;
Matches 269; Conservative 30; Mismatches 21; Indels 0; Gaps 0;

QY 5 RSGIAQVTELLIGKTPLVYLNKADGCVARVAAKLEMEPCSSVKDRIGFSMIDAEK 64
DB 4 RIGIANDVTELLIGNTPLVYLNVAEGCVGRVAAKLEMEPCSSVKDRIGFSMIDAEK 63
QY 65 ITPGKSVLIBPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILLAFGAEVLVLTDP 124
DB 64 IKPGESVLIBPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILLAFGAEVLVLTDP 123
QY 125 AKGMKGAIVAKAEIILAKTPNAYILQOFENPANPKHYETTGPETWKSDDKIDAFVSGIG 184
DB 124 AKGMKGAIVAKAEIILAKTPNAYILQOFENPANPKHYETTGPETWKSDDKIDAFVSGIG 183
QY 185 TGGTITGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEVNLLD 244
DB 184 TGGTITGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEVNLLD 243
QY 245 EVQOISSDEAIEITAKLALKEGLFVGISGAAAAAFOIAKRPENAGKLIYAVPSPFGERY 304
DB 244 EVQOISSDESIDMARLLAKREGLLVGISSGAAAAAIIKAKRPENAGKLIYAVPSPFGERY 303
QY 305 LSSVLFESVRRRAESMTFE 324
DB 304 LSTVLFDAITRKBAESMTFQ 323

RESULT 6
CYK2_ARATH STANDARD; PRT; 322 AA.
ID CYK2_ARATH STANDARD; PRT; 322 AA.
AC P47998; O23343; Q42570;
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cysteine synthase (EC 2.5.1.47) (O-acetylserine sulphydrylase) (O-
DE acetylserine (thiol)-lyase) (CSase A) (CS-A) (OAS-TL A) (Cys-3A)
DE (At.OAS.5-8).
GN OAS1 OR OAS1 OR AT4G14880 OR DL3480C.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_Taxid=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia; TISSUE=Leaf;
RX MEDLINE=94364481; PubMed=8082776;
RA Hell R., Bork C., Bogdanova N., Frolov I., Hauschild R.;
RT "Isolation and characterization of two cDNAs encoding for compartment
RT specific isoforms of O-acetylserine (thiol) lyase from Arabidopsis
RT thaliana."
RT FEBS Lett. 351:257-262 (1994).
RN [2]
RP REVISIONS.
RA Hell R.;
RP Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=95246850; PubMed=7729527;
RA Barroso C., Vega J.M., Gotor C.;
RT "A new member of the cytosolic O-acetylserine(thiol)lyase gene family
RT in Arabidopsis thaliana."
RN [4]
RX FEBS Lett. 363:1-5 (1995).

```

RA SEQUENCE FROM N.A.  
 RA STRAIN=cv. Columbia;  
 RA MEDLINE=20400348; PubMed=10940562;  
 RA Jost R., Berkowitz O., Wirtz M., Hopkins L., Hawkesford M.J., Helli R.;  
 RA "Genomic and functional characterization of the oas gene family  
 RA encoding O-acetylserine (thiol) lyases, enzymes catalyzing the final  
 RA step in cysteine biosynthesis in Arabidopsis thaliana.";  
 RA Gene 253:237-247(2000).  
 [5]  
 RA SEQUENCE FROM N.A.  
 RA STRAIN=cv. Columbia;  
 RA MEDLINE=9812113; PubMed=9461215;  
 RA Bevan M., Bancroft I., Bent E., Love K., Goodman H.M., Dean C.,  
 RA Bergkamp R., Dirkse W., van Staveren M., Stiekema W., Drost L.,  
 RA Ridley P., Hudson S.-A., Patel K., Murphy G., Piffanelli P.,  
 RA Wedler H., Wedler E., Wambutt R., Weitzenegger T., Pohl T., Terry N.,  
 RA Gielen J., Villarroel R., De Clercq R., van Montagu M., Lecharny A.,  
 RA Aubourg S., Gy I., Kreis M., Lao N., Kavanagh T., Herpel S.,  
 RA Kotter P., Entian K.-D., Rieger M., Schaefer M., Funk B.,  
 RA Mueller-Auer S., Silvey M., James R., Monfort A., Pons A.,  
 RA Puigdomenech P., Douka A., Vouklatou E., Milioni D., Hatzopoulos P.,  
 RA Piravandi B., Obermaier B., Hilbert H., Duesterhoeft A., Moores T.,  
 RA Jones J.D.G., Eneva T., Palme K., Benes V., Rechmann S., Ansoorge W.,  
 RA Cooke R., Berger C., Delseny M., Voet M., Volckaert G., Mewes H.-W.,  
 RA Klosterman S., Schueller C., Chalwatzis N.;  
 RA "Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of  
 RA Arabidopsis thaliana.";  
 RA Nature 391:485-488(1998).  
 [6]  
 RA SEQUENCE FROM N.A.  
 RA STRAIN=cv. Columbia;  
 RA MEDLINE=20083488; PubMed=10617198;  
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,  
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N.,  
 RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,  
 RA Weichselgartner M., de Simone V., Obermaier B., Maché R., Mueller M.,  
 RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schudthelini T.,  
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,  
 RA Vos P., Hohelsel J., Zimmermann W., Wedler H., Ridley P.,  
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,  
 RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,  
 RA Braeken M., Weltjens I., Voet M., Baetiaens I., Aert R., Defoor E.,  
 RA Weitzenegger T., Bothe G., Ransperger U., Hilbert H., Braun M.,  
 RA Holzner E., Brandt A., Peters S., van Staveren M., Dirkse W.,  
 RA Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,  
 RA Berner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,  
 RA De Keyser A., Buysaert C., Gielen J., Villarroel R., De Clercq R.,  
 RA Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,  
 RA Clark L., Doggett J., Hall S., Kay M., Leonard N., McIlroy K., Mayes R.,  
 RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,  
 RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,  
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fartmann B., Grandérath K., Dauner D., Herzi A.,  
 RA Neumann S., Argitlou A., Vitale D., Liguori R., Piravandi E.,  
 RA Massenet O., Quigley F., Clabaud G., Muendlein A., Felber R.,  
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,  
 RA Chefdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,  
 RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,  
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,  
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,  
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,  
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,  
 RA Parnell L., Dedha N., Gnoj L., Schutz K., Huang E., Spiegel L.,  
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,  
 RA Stoneking T., Kallicki J., Graves T., Harmon G., Edwards J.,  
 RA Latreille P., Courney L., Cloud J., Abbott A., Scott K., Johnson D.,  
 RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,  
 RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.W.,  
 RA Nelson J., Spieth J., Ryan B., Andrews S., Geisel C., Layman D.,  
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,  
 RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,  
 RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,  
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,

RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,  
 RA Chen B., Marra M.A., Martienssen R., McCombie W.R.;  
 RA "Sequence and analysis of chromosome 4 of the plant Arabidopsis  
 RA thaliana.";  
 RA Nature 402:769-777(1999).  
 RA Nature 402:769-777(1999).  
 CC -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +  
 CC acetate.  
 CC -1- COFACTOR: Pyridoxal phosphate.  
 CC -1- PATHWAY: Cysteine biosynthesis.  
 CC -1- SUBUNIT: Homodimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-  
 CC synthase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X80376; CAA56593.2; -;  
 DR EMBL; X84097; CAA58893.1; -;  
 DR EMBL; AJ272027; CAB72932.1; -;  
 DR EMBL; Z97337; CAB10267.1; -;  
 DR EMBL; AL161540; CAB78530.1; -;  
 DR PIR; A71412; A71412.  
 DR HSSP; P12674; 10AS.  
 DR InterPro; IPR001926; B6\_enzyme\_beta.  
 DR InterPro; IPR001216; Cys\_synthase\_BS.  
 DR InterPro; IPR005859; Cys\_synthase.  
 DR InterPro; IPR005856; Cys\_synthase.  
 DR Pfam; PF00291; PALP; 1.  
 DR TIGRPFAM; TIGR01139; cySK; 1.  
 DR TIGRPFAM; TIGR01136; cySKM; 1.  
 DR PROSITE; PS00901; CYS\_SYNTHASE; 1.  
 DR TRANSFERASE; Cysteine biosynthesis; Pyridoxal phosphate.  
 DR BINDING; 46 46 PYRIDOXAL PHOSPHATE (By SIMILARITY).  
 FT DOMAIN 272 277 POLY-ALA.  
 FT CONFLICT 273 273 A -> E (in REF. 3).  
 FT SEQUENCE 322 AA; 33805 MW; 5B3E7F3D9DA5908B CRC64;  
 SQ  
 Query Match 85.3%; Score 1384; DB 1; Length 322;  
 Best Local Similarity 83.4%; Pred. No. 1.1e-87;  
 Matches 266; Conservative 30; Mismatches 23; Indels 0; Gaps 0;  
 Oy 6 SGIAKDVTELTIGTPTLVYLNKLDGCVARVAKLELWPCSSVYKDRIGYSMIADAERKGL 65  
 Db 3 SRIAKDVTELTIGTPTLVYLNKLDGCVARVAKLELWPCSSVYKDRIGYSMIADAERKGL 62  
 Oy 66 ITPGKSVLIEPTSGNTGIGLAFMAARAGYKLIITPMASLSRRRIILAFGAELVLTDP 125  
 Db 63 IKPESVLIIEPTSGNTGIGLAFMAARAGYKLIITPMASLSRRRIILAFGAELVLTDP 122  
 Oy 126 KAKGAVQKAEIILAKTPNAVILQEPENPANPKVHYETTGPEIWKSGDKIDAFVSGIGT 185  
 Db 123 KAKGAVQKAEIILAKTPNAVILQEPENPANPKVHYETTGPEIWKSGDKIDAFVSGIGT 182  
 Oy 186 GGTITGAGKYLKEQNPNIKIGVEPVESPVLSGKPGPHKIIGIGAGFIPGVLBNVLDE 245  
 Db 183 GGTITGAGKYLKEQNPNIKIGVEPVESPVLSGKPGPHKIIGIGAGFIPGVLBNVLDE 242  
 Oy 246 VVOISSDEAIEITAKLILKEGLFVGISSGAAAFQIAKRPENAGKLIIVAVPSGGERY 305  
 Db 243 VVOISSDEAIEITAKLILKEGLFVGISSGAAAFQIAKRPENAGKLIIVAVPSGGERY 302  
 Oy 306 LSVLFESVRRRAESMTPE 324  
 Db 303 LSVLFESVRRRAESMTPE 321  
 RESULT 7



CYSK MAIZE  
ID CYSK MAIZE STANDARD; PRT; 325 AA.  
AC P80608;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cysteine synthase (EC 2.5.1.47) (O-acetylserine sulphydrylase)  
DE (O-acetylserine (Thiol)-lyase) (CSase) (OAS-TL).  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. B73; TISSUE=Root;  
RA Brander K.A., Owttrim G.W., Brumold C.,  
RT "Isolation of a cDNA encoding a putative chloroplastic isoform of  
RL Cysteine synthase from maize."  
RL (in) Plant Gene Register PGR95-031.  
RN [2]  
RP SEQUENCE OF 11-25.  
RC TISSUE=Coleoptile;  
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,  
RA Bernollet J.-C., Zivy M., de Vienne D.,  
RT "The maize two dimensional gel protein database: towards an integrated  
RL genome analysis program."  
RL Theor. Appl. Genet. 93:997-1005(1996).  
CC -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +  
CC acetate.  
CC -1- COFACTOR: Pyridoxal phosphate (By similarity).  
CC -1- PATHWAY: Cysteine biosynthesis.  
CC -1- SUBUNIT: Homodimer (By similarity).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-  
CC synthase family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X85803; CA59798.1; -  
DR PIR; S52738; S52738.  
DR Maize-2DPAGE; P80608; COLBOPTILE.  
DR HSSP; P12674; 1FCJ.  
DR MaizeDB; 123922; -  
DR InterPro; IPR001926; B6\_enzyme\_beta.  
DR InterPro; IPR001216; Cys\_synthase\_BS.  
DR InterPro; IPR005859; Cys\_synthk.  
DR InterPro; IPR005856; Cys\_synthKM.  
DR Pfam; PF00291; PALP; 1.  
DR TIGRfams; TIGR01139; cysK; 1.  
DR TIGRfams; TIGR01136; cysKM; 1.  
DR PROSITE; PS00901; CYS\_SYNTHASE; 1.  
KW Transferase; Cysteine biosynthesis; Pyridoxal phosphate.  
FT BINDING 49 49 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
SQ SEQUENCE 325 AA; 34206 MW; 3213A69326D7CEBD CRC64;  
Query Match 84.9%; Score 1378; DB 1; Length 325;  
Best Local Similarity 84.3%; Pred. No. 2.8e-87;  
Matches 268; Conservative 22; Mismatches 28; Indels 0; Gaps 0;  
QY 8 IAKDVTLLIGTPTLVYLNKLADGCVARVAKLEMEPCSSVYKDRIGYSMTADAEKGLIT 67  
DB 8 IAKDVTLLIGTPTLVYLNKLADGCVARVAKLEMEPCSSVYKDRIGYSMTADAEKGLIT 67  
QY 68 PGKSVLIBPTSGNTGIGLAFFMAAARGYKLTITMFASSLSERRIILLAFGAELVLTDPK 127  
FT BINDING 49 49 PYRIDOXAL PHOSPHATE (BY SIMILARITY).

DB 68 PGVSVLIBPTSGNTGIGLAFFMAAARGYKLTITMFASSLSERRIILLAFGAELVLTDP 127  
QY 128 MGAIVOKAEIILAKTPNAVYILQFENPANPKVHYETTGPEIMWKS DGKIDAFVSGIGTGG 187  
DB 128 MGAIVOKAEIILAKTPNAVYILQFENPANPKVHYETTGPEIMWKS DGKIDAFVSGIGTGG 187  
QY 188 TITGAGKYLKEQNPNKILGVBPVPSVYLSGKRGDPHKIQIGIGAGFTPGVLEVNILDEVY 247  
DB 188 TITGAGKYLKEQNPNKILGVBPVPSVYLSGKRGDPHKIQIGIGAGFTPGVLEVNILDEVY 247  
QY 248 QISSDEAIEFAKLALKEGLFVGISSGAAAAAFAQIAKRPENAGKLTIVAVPSPGGRYLS 307  
DB 248 QISSDEAIEFAKLALKEGLFVGISSGAAAAAFAQIAKRPENAGKLTIVAVPSPGGRYLS 307  
QY 308 SVLPESVVRRAESMTPEP 325  
DB 308 SVLPESVVRRAESMTPEP 325  
QY 308 SVLPESVVRRAESMTPEP 325  
DB 308 SVLPESVVRRAESMTPEP 325  
RESULT 8  
CYK2\_ORYSA  
ID CYK2\_ORYSA STANDARD; PRT; 325 AA.  
AC Q9XEA8;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cysteine synthase (EC 2.5.1.47) (O-acetylserine sulphydrylase)  
DE (O-acetylserine (Thiol)-lyase) (CSase) (OAS-TL).  
GN RCS3.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzeae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Nipponbare;  
RX MEDLINE=99196994; Pubmed=10095115;  
RA Nakamura T., Yamaguchi Y., Sano H.,  
RT "Four rice genes encoding cysteine synthase: isolation and  
RL differential responses to sulfur, nitrogen and light."  
RL Gene 229:155-161(1999).  
CC -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +  
CC acetate.  
CC -1- COFACTOR: Pyridoxal phosphate (By similarity).  
CC -1- PATHWAY: Cysteine biosynthesis.  
CC -1- SUBUNIT: Homodimer (By similarity).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-  
CC synthase family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF073697; AAD23909.1; -  
DR HSSP; P12674; 1FCJ.  
DR AND-2DPAGE; Q9XEA8; -  
DR Gramene; Q9XEA8; -  
DR InterPro; IPR001926; B6\_enzyme\_beta.  
DR InterPro; IPR001216; Cys\_synthase\_BS.  
DR InterPro; IPR005859; Cys\_synthk.  
DR InterPro; IPR005856; Cys\_synthKM.  
DR Pfam; PF00291; PALP; 1.  
DR TIGRfams; TIGR01139; cysK; 1.  
DR TIGRfams; TIGR01136; cysKM; 1.  
DR PROSITE; PS00901; CYS\_SYNTHASE; 1.  
KW Transferase; Cysteine biosynthesis; Pyridoxal phosphate.  
FT BINDING 49 49 PYRIDOXAL PHOSPHATE (BY SIMILARITY).

FT DOMAIN 275 280 POLY-ALA.  
SQ SEQUENCE 325 AA; 34306 MW; 8F6B0CDCCD7DCAC4 CRC64;  
Query Match 84.8%; Score 1377; DB 1; Length 325;  
Best Local Similarity 80.9%; Pred. No. 3.3e-87;  
Matches 263; Conservative 28; Mismatches 34; Indels 0; Gaps 0;

QY 1 MAVERSGIAKDVTELGKTPLYVNLKADGCVARVAAKLEIMBPCSSVKDRIGYSMTADA 60  
DB 1 MAESGSIASDVTAALIGNTPLVYLNKVDGCEAQIAKLEIMBPCSSVKDRIGYSMTADA 60  
QY 61 EEKGLTPGKSVLIEPTSGNTGIGLAFAAARGYKLIITMPASMSLERIILLAFGAELV 120  
DB 61 EEKGLTPGKSVLIEPTSGNTGIGLAFAAARGYKLIITMPASMSMERIILLAFGAELV 120  
QY 121 LTDPKAKGAVOKABEILAKTPNAYTLQGFENPANPKVHYETTGPETWKSDEKIDAFV 180  
DB 121 LTDPKAKGAVOKABEILAKTPNAYTLQGFENPANPKVHYETTGPETWKSDEKIDAFV 180  
QY 181 SGIGTGTTGAGKYLKEQNPNIKIGVEPVESPVLSGKPGPKIQIGAGFIPGVLEV 240  
DB 181 SGIGTGTTGAGKYLKEQNPNIKIGVEPVESPVLSGKPGPKIQIGAGFIPGVLEV 240  
QY 241 NLDEVVQISSDEAIEITAKLALKEGLFVGISSGAAAAAFAQIAKRPENAGKLIIVAVPES 300  
DB 241 NLDEVVQISSDEAIEITAKLALKEGLFVGISSGAAAAAFAQIAKRPENAGKLIIVAVPES 300  
QY 301 FGERYLSSVLFESVRRREASMTFEP 325  
DB 301 FGERYLSSVLFESVRRREASMTFEP 325

RESULT 9  
CYSK WHEAT  
ID CYSK\_WHEAT STANDARD; PRT; 325 AA.  
AC P38076;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cysteine synthase (EC 2.5.1.47) (O-acetylserine sulfinhydrilase)  
DN (O-acetylserine (Thiol)-lyase) (CSase A) (OAS-TL A).  
OS Triticum aestivum (Wheat).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;  
OC Triticaceae; Triticum.  
OX NCBI\_TaxID=4565;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Leaf;  
RX MEDLINE=94100990; PubMed=8275097;  
RA Youssefian S., Nakamura M., Sano H.;  
RT "Tobacco plants transformed with the O-acetylserine (thiol) lyase  
RT gene of wheat are resistant to toxic levels of hydrogen sulphide  
RT gas.";  
RL Plant J. 4:759-769(1993).  
CC -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +  
CC acetate.  
CC -1- COFACTOR: Pyridoxal phosphate.  
CC -1- PATHWAY: Cysteine biosynthesis.  
CC -1- SUBUNIT: Homodimer (By similarity).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-  
CC synthase family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; D13153; BAA02438.1; -.  
DR PIR; J50762; J50762.  
DR HSSP; P12674; 1OAS.  
DR InterPro; IPR001926; B6\_enzyme\_beta.  
DR InterPro; IPR001216; Cys\_synthase\_BS.  
DR InterPro; IPR005859; Cys\_synthase.  
DR InterPro; IPR005856; Cys\_synthase.  
DR Pfam; PF00291; PALP; 1.  
DR TIGRfam; TIGR01139; CysK; 1.  
DR TIGRfam; TIGR01136; CysK; 1.  
DR PROSITE; PS00901; CYS\_SYNTHASE; 1.  
DR Transferase; Cysteine biosynthesis; Pyridoxal phosphate.  
FT BINDING 50 PYRIDOXAL PHOSPHATE (By similarity).  
FT DOMAIN 50 POLY-ALA.  
SQ SEQUENCE 325 AA; 34114 MW; 1CEFBF29F341DCA CRC64;  
Query Match 84.0%; Score 1364; DB 1; Length 325;  
Best Local Similarity 82.3%; Pred. No. 2.5e-86;  
Matches 261; Conservative 27; Mismatches 29; Indels 0; Gaps 0;

QY 8 IAKDVTELGKTPLYVNLKADGCVARVAAKLEIMBPCSSVKDRIGYSMTADAERGLIT 67  
DB 9 IAKDVTELGKTPLYVNLKADGCVARVAAKLEIMBPCSSVKDRIGYSMTADAERGLIT 67  
QY 68 PGKSVLIEPTSGNTGIGLAFAAARGYKLIITMPASMSLERIILLAFGAELVITDPAG 127  
DB 69 PGKSVLIEPTSGNTGIGLAFAAARGYKLIITMPASMSMERIILLAFGAELVITDPAG 127  
QY 128 MKGAVOKABEILAKTPNAYTLQGFENPANPKVHYETTGPETWKSDEKIDAFVSGIGTG 187  
DB 129 MKGAVOKABEILAKTPNAYTLQGFENPANPKVHYETTGPETWKSDEKIDAFVSGIGTG 187  
QY 188 TITGAGKYLKEQNPNIKIGVEPVESPVLSGKPGPKIQIGAGFIPGVLEVNLDEVV 247  
DB 189 TITGAGKYLKEQNPNIKIGVEPVESPVLSGKPGPKIQIGAGFIPGVLEVNLDEVV 247  
QY 248 QISSDEAIEITAKLALKEGLFVGISSGAAAAAFAQIAKRPENAGKLIIVAVPESGERYLS 307  
DB 249 QISSDEAIEITAKLALKEGLFVGISSGAAAAAFAQIAKRPENAGKLIIVAVPESGERYLS 307  
QY 308 SVLFESVRRREASMTFEP 324  
DB 309 SVLFESVRRREASMTFEP 324

RESULT 10  
CYKL ORYSA  
ID CYKL\_ORYSA STANDARD; PRT; 321 AA.  
AC Q9XBA6;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cysteine synthase (EC 2.5.1.47) (O-acetylserine sulfinhydrilase)  
DN (O-acetylserine (Thiol)-lyase) (CSase) (OAS-TL).  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Nipponbare;  
RX MEDLINE=99196994; PubMed=10095115;  
RA Nakamura T., Yamaguchi Y., Sano H.;  
RT "Four rice genes encoding cysteine synthase: isolation and  
RT differential responses to sulfur, nitrogen and light.";  
RL Gene 229:155-161(1999).  
CC -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +  
CC acetate.  
CC -1- COFACTOR: Pyridoxal phosphate (By similarity).  
CC -1- PATHWAY: Cysteine biosynthesis.  
CC -1- SUBUNIT: Homodimer (By similarity).  
CC -----

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-synthase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF073695; AAD23907.1; -.  
 CC HSSP; P12674; IFCU.  
 CC Gramene; Q9XEA6; -.  
 CC InterPro; IPR001926; B6\_enzyme\_beta.  
 CC InterPro; IPR001216; Cys\_synthase\_BS.  
 CC InterPro; IPR005859; Cys\_synthk.  
 CC InterPro; IPR005856; Cys\_synthk.  
 CC Pfam; PF00291; PALP; 1.  
 CC TIGRFAMS; TIGR01139; cysK; 1.  
 CC TIGRFAMS; TIGR01136; cysK; 1.  
 CC PROSITE; PS00901; CYS\_SYNTHASE; 1.  
 CC Transferrase; Cysteine biosynthesis; Pyridoxal phosphate.  
 CC BINDING 46 46 PYRIDOXAL PHOSPHATE (By SIMILARITY).  
 CC DOMAIN 272 277 POLY-ALA.  
 CC SEQUENCE 321 AA; 33868 MW; 9DCC2A2C9A6D63A6 CRC64;  
 QY 8 IAKDVTLEIKTLPVLYLNKLADGCVARVAAKLEMEPCSSVKDRIGYSMTADAEKGLIT 67  
 Db 5 IAKDVTLEIKTLPVLYLNKLADGCVARVAAKLEMEPCSSVKDRIGYSMTADAEKGLIT 64  
 QY 68 PGKSVLIEPTSGNTGIGLAFMAAARGYKLTITMPASMSLERRIILAFGAELVLTDPAG 127  
 Db 65 PGKSVLIEPTSGNTGIGLAFMAAARGYKLTITMPASMSLERRIILAFGAELVLTDPAG 124  
 QY 128 MKGAVOKAEELIATNPAYILLOQFENPANPKVHYETTGPRIWKGSDGKIDAFVSGIGTGG 187  
 Db 125 MKGAVOKAEELIATNPAYILLOQFENPANPKVHYETTGPRIWKGSDGKIDAFVSGIGTGG 184  
 QY 188 TITGAGKYLKEONPNIKLIGVEPVESPVLSGGKPGPHKIQIGAGFIPGVLEVNLDDEV 247  
 Db 185 TITGAGKYLKEONPNIKLIGVEPVESPVLSGGKPGPHKIQIGAGFIPGVLEVNLDDEV 244  
 QY 248 QISSDEAIEIAKLAEKGLFVGISGGAFAAFAQIAKRPENAGKLIIVAVPSPGERYLS 307  
 Db 245 QISSDEAIEIAKLAEKGLFVGISGGAFAAFAQIAKRPENAGKLIIVAVPSPGERYLS 304  
 QY 308 SVLFESVRRREASMTTFE 324  
 Db 305 SVLFESVRRREASMTTFE 321  
 RESULT 11  
 CYSL\_SOLTU STANDARD; PRT; 386 AA.  
 AC 081155;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cysteine synthase, chloroplast precursor (BC 2.5.1.47) (O-acetylserine  
 DE sulphydrylase) (O-acetylserine (Thiol)-lyase) (CSase B) (CS-B) (OAS-TL  
 DB B).  
 OS Solanum tuberosum (Potato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatoxophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Lamiales; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4113;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Berolina;  
 RA Heese H.; Hoefgen R.;  
 RT "Isolation of cDNAs encoding cytosolic and plastidic cysteine synthase  
 RL isoforms from Solanum tuberosum."  
 CC (in) Plant Gene Register PGR98-057.  
 CC -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +  
 CC acetate.  
 CC -1- COFACTOR: Pyridoxal phosphate (By similarity).  
 CC -1- PATHWAY: Cysteine biosynthesis.  
 CC -1- SUBUNIT: Homodimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Chloroplast stroma; chromoplast (By  
 CC similarity).  
 CC -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-  
 CC synthase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF044173; AAC25636.1; -.  
 CC PIR; T07002; T07002.  
 CC HSSP; P12674; IOAS.  
 CC InterPro; IPR001926; B6\_enzyme\_beta.  
 CC InterPro; IPR001216; Cys\_synthase\_BS.  
 CC InterPro; IPR005859; Cys\_synthk.  
 CC InterPro; IPR005856; Cys\_synthk.  
 CC Pfam; PF00291; PALP; 1.  
 CC TIGRFAMS; TIGR01139; cysK; 1.  
 CC TIGRFAMS; TIGR01136; cysK; 1.  
 CC PROSITE; PS00901; CYS\_SYNTHASE; 1.  
 CC Transferrase; Cysteine biosynthesis; Pyridoxal phosphate; Chloroplast;  
 CC TRANSIT peptide.  
 CC TRANSIT 1 52 CHLOROPLAST (By SIMILARITY).  
 CC CHAIN 53 386 CYSTEINE SYNTHASE.  
 CC FT BINDING 110 110 PYRIDOXAL PHOSPHATE (By SIMILARITY).  
 CC SEQUENCE 386 AA; 41044 MW; 8EB3FEB2084D9594 CRC64;  
 Query Match 81.0%; Score 1314; DB 1; Length 386;  
 Best Local Similarity 78.1%; Pred. No. 8e-83;  
 Matches 249; Conservative 38; Mismatches 32; Indels 0; Gaps 0;  
 QY 3 VERSGIADVTLEIKTLPVLYLNKLADGCVARVAAKLEMEPCSSVKDRIGYSMTADAE 62  
 Db 64 IEGNIADVTQLIGNTPMVLYNTIAKGVANIAAKLEINEPCSSVKDRIGYSMTADAE 123  
 QY 63 KGLTPKSVLIEPTSGNTGIGLAFMAAARGYKLTITMPASMSLERRIILAFGAELVLT 122  
 Db 124 KGLISPGKTVLVEPTSGNTGIGLAFIAASRGYKLTITMPASMSLERRIILAFGAELVLT 183  
 QY 123 DPAKMGKAVOKAEELIATNPAYILLOQFENPANPKVHYETTGPRIWKGSDGKIDAFVSG 182  
 Db 184 DPAKMGKAVOKAEELIATNPAYILLOQFENPANPKVHYETTGPRIWKGSDGKIDAFVSG 243  
 QY 183 IGTGTTTGAAGKYLKEONPNIKLIGVEPVESPVLSGGKPGPHKIQIGAGFIPGVLEVN 242  
 Db 244 IGTGTTTGAAGKYLKEONPNIKLIGVEPVESPVLSGGKPGPHKIQIGAGFIPGVLEVN 303  
 QY 243 LDEVVOISSDEAIEIAKLAEKGLFVGISGGAFAAFAQIAKRPENAGKLIIVAVPSPFG 302  
 Db 304 MDEVVIEISSDEAIEIAKLAEKGLFVGISGGAFAAFAQIAKRPENAGKLIIVAVPSPFG 363  
 QY 303 ERYLGSVLFESVRRREASMTTFE 321  
 Db 364 ERYLGSVLFESVRRREASMTTFE 382  
 RESULT 12  
 CYSL\_SPIOL



ID CYS1 SPIOL STANDARD; PRT; 383 AA.  
 AC P32260; Q33137;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cysteine synthase, chloroplast precursor (EC 2.5.1.47) (O-acetylserine  
 sulfhydrylase) (O-acetylserine (Thiol)-lyase) (CSase B) (CS-B) (OAS-TL  
 B).  
 GN CYSK.  
 OS Spinacia oleracea (Spinach).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllales; Amaranthaceae; Spinacia.  
 OC NCBI\_TaxID=3562;  
 RN [1]  
 RP MEDLINE=94009590; PubMed=8405359;  
 RA Saito K., Tatsuguchi K., Murakoshi I., Hirano H.;  
 RT "cDNA cloning and expression of cysteine synthase B localized in  
 chloroplasts of Spinacia oleracea.";  
 RL FEBS Lett. 324:247-252(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=leaf;  
 RX MEDLINE=93143317; PubMed=8424655;  
 RA Rolland N., Droux M., Lebrun M., Douce R.;  
 RT "O-acetylserine(thiol) lyase from spinach (Spinacia oleracea L.) leaf:  
 cDNA cloning, characterization, and overexpression in Escherichia  
 coli of the chloroplast isoform.";  
 RL Arch. Biochem. Biophys. 300:213-222(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Marathon; TISSUE=leaf;  
 RX MEDLINE=94105326; PubMed=8278530;  
 RA Helli R., Schuster G., Grisse W.;  
 RT "An O-acetylserine (thiol) lyase cDNA from spinach.";  
 RL Plant Physiol. 102:1057-1058(1993).  
 RN [4]  
 RP DISCUSSION OF SEQUENCE.  
 RX MEDLINE=93356752; PubMed=7916619;  
 RA Rolland N., Job D., Douce R.;  
 RT "Common sequence motifs coding for higher-plant and prokaryotic O-  
 acetylserine (thiol)-lyases: bacterial origin of a chloroplast  
 transit peptide?";  
 RL Biochem. J. 293:829-833(1993).  
 RN [5]  
 RP PARTIAL SEQUENCE, AND CHARACTERIZATION.  
 RX MEDLINE=96184908; PubMed=8617276;  
 RA Rolland N., Ruffet M.-L., Job D., Douce R., Droux M.;  
 RT "Spinach chloroplast O-acetylserine (thiol)-lyase exhibits two  
 catalytically non-equivalent pyridoxal-5'-phosphate-containing active  
 sites.";  
 RL Eur. J. Biochem. 236:272-282(1996).  
 CC -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +  
 acetate.  
 CC -1- COFACTOR: Pyridoxal phosphate.  
 CC -1- PATHWAY: Cysteine biosynthesis.  
 CC -1- SUBUNIT: Homodimer.  
 CC -1- SUBCELLULAR LOCATION: Chloroplast stroma; chromoplast.  
 CC -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-  
 synthase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; D14722; BAA03542.1; -;  
 CC EMBL; X66860; CAA47329.1; -;  
 CC EMBL; L05184; AAA16973.1; -;  
 CC RN

DR PIR; S29733; S29733.  
 DR PIR; T09000; T09000.  
 DR HSSP; P12674; 1FCJ.  
 DR InterPro; IPR001926; B6\_enzyme\_beta.  
 DR InterPro; IPR001216; Cys\_synthase\_BS.  
 DR InterPro; IPR005859; Cys\_synthK.  
 DR InterPro; IPR005856; Cys\_synthKM.  
 DR Pfam; PF00291; PALP; 1.  
 DR TIGRFAMs; TIGR01339; cysK; 1.  
 DR TIGRFAMs; TIGR01336; cysKM; 1.  
 DR PROSITE; PS00901; CYS\_SYNTHASE; 1.  
 KM Transferase; Cysteine biosynthesis; Pyridoxal phosphate; Chloroplast;  
 KM Transit peptide.  
 FT TRANSIT 1 52 CHLOROPLAST.  
 FT CHAIN 53 383 CYSTEINE SYNTHASE.  
 FT BINDING 108 108 PYRIDOXAL PHOSPHATE.  
 FT DOMAIN 334 339 POLY-ALA.  
 FT CONFLICT 12 12 L -> I (IN REF. 1).  
 FT CONFLICT 166 166 E -> EKESYLE (IN REF. 3).  
 FT CONFLICT 252 258 GRYLKER -> DGTSKNA (IN REF. 3).  
 FT CONFLICT 274 280 ILGGKP -> YPLVESA (IN REF. 3).  
 FT CONFLICT 334 335 AA -> RR (IN REF. 2).  
 FT CONFLICT 334 335 AA -> RG (IN REF. 3).  
 FT CONFLICT 379 380 NM -> KL (IN REF. 3).  
 FT CONFLICT 383 383 E -> EI (IN REF. 3).  
 SQ SEQUENCE 383 AA; 40637 MW; CE3725796D3A3F93 CRC64;  
 Query Match 78.3%; Score 1271; DB 1; Length 383;  
 Best Local Similarity 73.4%; Pred. No. 6.9e-80;  
 Matches 234; Conservative 52; Mismatches 33; Indels 0; Gaps 0;  
 QY 3 VERSGIADVTLLIGKTPLVYLNKLADGCVARVAAKLEIMPCSSVDRIQYSMIADAE 62  
 DB 62 EKGLEIAEDVSQIIGKTPMVTNNVSKSVANIAKLESMEPCCSVDRIQYSMIDAEQ 121  
 QY 63 KGLITPKSVLIBPISGNTGIGLAFMAAARGYKLIITPMASMSLERIILAFGAELVLT 122  
 DB 122 KGVITPKKTLVEPTSGNTGIGLAFIAARGYKITLTMPASMSMERVILKAFGAELVLT 181  
 QY 123 DPAKMGGAVOAEELIAPKTPNAVYILQGFENPANKVHYETTGPEIWKSDGKIDAFVSG 182  
 DB 182 DPAKMGGAVERAEELIKTPDPSYMLQGFDPNPNPKIHYETTGPEIWKSDGKIDAFVAG 241  
 QY 183 IGTGTTGAGTYLKEQNPNIKLGVEPVESPVLGGKPGPHKIQGIGAGFIPGVLEVNL 242  
 DB 242 IGTGTTSGVGRYTKERNPGVQVIGIEPESNILLSGKPGPHKIQGIGAGFIPGVLEVNL 301  
 QY 243 IDEVVOISSDEAIEETAKLLALKEGLFVGISGAAAAAAAFQIAKRPENAGELIVAPPSFG 302  
 DB 302 MDEVIVSSBEAVEMAKQELVGVISGAAAAAAAVRIGKRPENAGELIVAPPSFG 361  
 QY 303 ERYLSSVLFESVREAEBSM 321  
 DB 362 ERYLSSVLFESVREAEBSM 380  
 RESULT 13  
 CYSM\_ARATH STANDARD; PRT; 424 AA.  
 ID CYSM\_ARATH  
 AC Q43725; Q9M128; Q9M440;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cysteine synthase, mitochondrial precursor (EC 2.5.1.47) (O-  
 acetylserine sulfhydrylase) (O-acetylserine (Thiol)-lyase) (CSase C)  
 DE (CS-C) (OAS-TL C) (ACCS-C).  
 GN OASC OR ACS 1 OR AT3G59760 OR F24G16.30.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC Eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OK NCBI\_TaxID=3702;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia; TISSUE=flower;  
 RX MEDLINE=99252742; PubMed=10319184;  
 RA Hesse H., Lipke J., Altmann T., Hofgen R.;  
 RT "Molecular cloning and expression analyses of mitochondrial and  
 RT plastidic isoforms of cysteine synthase (O-acetylserine(thiol) lyase)  
 RT from Arabidopsis thaliana";  
 RL Amino Acids 16:113-131(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016720; PubMed=11130713;  
 RA Salanoubat M., Lemcke K., Rieger M., Ansoorge W., Unseld M.,  
 RA Fartmann B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,  
 RA Delseny M., Boutry M., Grivell L.A., Maché R., Puigdomenech P.,  
 RA de Simone V., Choisme N., Artiguenave F., Robert C., Brothier P.,  
 RA Wincker P., Cattolico L., Weissenbach J., Saurin W., Quétier F.,  
 RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,  
 RA Wurmback E., Drzonek H., Erfle H., Jordan N., Bangert S.,  
 RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,  
 RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simonati B.,  
 RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordstlex G.,  
 RA Reichelt J., Schärfe M., Schoen O., Barges M., Terol J., Clément J.,  
 RA Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,  
 RA Cooke R., Laudie M., Berger-Hauro C., Purnelle B., Masny D.,  
 RA de Haan M., Maarse A.C., Alcaraz J.-P., Cortet A., Casacuberta B.,  
 RA Monfort A., Argiriou A., Flores M., Liguori R., Vitale D.,  
 RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,  
 RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,  
 RA Rooney T., Rizzo M., Walte A., Utterback T., Fujii C.Y., Shea T.P.,  
 RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,  
 RA Pai G., Miltscher J., Sellers P., Gill J.E., Feldblyum T.V.,  
 RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,  
 RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,  
 RA Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,  
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,  
 RA Watanabe A., Yamada M., Yasuda M., Tabata S.;  
 RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis  
 RT thaliana";  
 RL Nature 408:820-822(2000).  
 RN [3]  
 RP SEQUENCE OF 38-424 FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=20400348; PubMed=10940562;  
 RA Jost R., Berkowitz O., Wirtz M., Hopkins L., Hawkesford M.J., Hell R.;  
 RT "Genomic and functional characterization of the cys gene family  
 RT encoding O-acetylserine (thiol) lyases, enzymes catalyzing the final  
 RT step in cysteine biosynthesis in Arabidopsis thaliana";  
 RL Gene 253:237-247(2000).  
 CC -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +  
 CC acetate.  
 CC -1- COFACTOR: Pyridoxal phosphate.  
 CC -1- PATHWAY: Cysteine biosynthesis.  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.  
 CC -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-  
 CC synthase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X81973; CA57498.1; -  
 DR EMBL; AL138647; CAB5795.1; -  
 DR EMBL; AJ271727; CAB71290.1; -  
 DR PIR; T52650; T52650.  
 DR HSSP; P12674; 1OAS.  
 DR InterPro; IPR001926; B6\_enzyme\_beta.  
 DR InterPro; IPR001216; Cys\_synthase\_BS.

DR InterPro; IPR005859; Cys\_synthK.  
 DR InterPro; IPR005856; Cys\_synthKM.  
 DR Pfam; PF00291; PALP; 1.  
 DR TIGRFAMS; TIGR01139; cysK; 1.  
 DR TIGRFAMS; TIGR01136; cysKM; 1.  
 DR PROSITE; PS00901; CYS\_SYNTHASE; 1.  
 KW Transferrase; Cysteine biosynthesis; Pyridoxal phosphate;  
 KW Mitochondrion; Transic peptide.  
 FT TRANSIT 1 ?  
 FT CHAIN 1 424  
 FT BINDING 148 148  
 FT DOMAIN 44 48  
 FT DOMAIN 374 379  
 FT CONFLICT 48 48  
 FT CONFLICT 67 67  
 FT CONFLICT 189 189  
 FT CONFLICT 308 308  
 FT CONFLICT 350 351  
 FT CONFLICT 379 379  
 FT CONFLICT 395 395  
 FT CONFLICT 403 403  
 FT CONFLICT 418 420  
 SO SEQUENCE 424 AA; 45182 MW; C37B5C04BF30BF2F CRC64;  
 Query Match 78.3%; Score 1271; DB 1; Length 424;  
 Best Local Similarity 73.9%; Pred. No. 7.7e-80;  
 Matches 241; Conservative 43; Mismatches 36; Indels 6; Gaps 1;  
 QY 2 AVERS-----GIADVTBLIGKTPVLYLNKLADGCVARVAATLMEPCSSVKRIGYS 55  
 DB 95 AVKRETPGDLINADNVSQLIGTPVLYNSIAKGVANIAALEIMEPCSSVKRIGYS 154  
 QY 56 MIADAEKGLITPGKSVLIBPESGNTGICLAFMAARGYKLTITPMASMSLERRILLAF 115  
 DB 155 MWTDAEQKGTISPCKSVLVEPTSGNTGICLAFIAASRGYRLITPMASMSMERVLKAF 214  
 QY 116 GAEVLVTPAKGMKGAIVOKAEIILAKTPNAYILQOFENPANPKHYETTPETIMKSDGK 175  
 DB 215 GAEVLVTPAKGMTGAVOKAEIILAKTPNAYILQOFENPANPKHYETTPETIMKSDGK 274  
 QY 176 IDAFVSGIGTGITITGAKYLKQNPINKLIGVEPVESPVLSGKQDPKHIQIGAGFIP 235  
 DB 275 VDIIVAGIGTGITITGVGRPIKKNPKTOVIGVEPTSDILSGKQDPKHIQIGAGFIP 334  
 QY 236 GYLEVNILDEVVOISSDEAIFAKLALKEGLFVGISSGAAAAAFOIAKRPENAKGLIV 295  
 DB 335 KNLQKIMDEVIAISSBEAIFAKLALKEGLVIGISSGAAAAAIVAKRPENAKGLIA 394  
 QY 296 AVPPSFGERYLSSVLPBSVRREAESM 321  
 DB 395 VVPPSFGERYLSTPLFQSIREEVERKM 420  
 RESULT 14  
 CYSYL\_ARATH STANDARD; PRT; 392 AA.  
 ID CYSYL\_ARATH  
 AC P47999; O22828; Q42568;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cysteine synthase, chloroplast precursor (EC 2.5.1.47) (O-acetylserine  
 DE sulfhydrylase) (O-acetylserine (thiol)-lyase) (CSase B) (CS-B) (OAS-TL  
 DE B) (ATCS-B) (cpACS1) (At.OAS-7-4).  
 GN OASB OR AT2G43750 OR F18019.14.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia; TISSUE=leaf;  
 RX MEDLINE=94364481; PubMed=8082776;

```

RA RA Hell R., Bork C., Bogdanova N., Frolov I., Hauschild R.;
RT RT "Isolation and characterization of two cDNAs encoding for compartment
RT RT specific isoforms of O-acetylserine (thiol) lyase from Arabidopsis
RT RT thaliana.";
RL RL FEBS Lett. 351:257-262(1994).
RM RM [2]
RN RN REVISIONS.
RA RA Hell R.;
RL RL Submitted (XXX-2000) to the EMBL/GenBank/DBJ databases.
RM RM [3]
RP RP SEQUENCE FROM N.A.
RC RC STRAIN=cv. Columbia;
RX RX MEDLINE=95334513; PubMed=7610184;
RA RA Hesse H., Altmann T.;
RT RT "Molecular cloning of a cysteine synthase cDNA from Arabidopsis
RL RL thaliana.";
RM RM Plant Physiol. 108:851-852(1995).
RN RN [4]
RP RP SEQUENCE FROM N.A.
RC RC STRAIN=cv. Columbia;
RX RX MEDLINE=20400348; PubMed=10940562;
RA RA Jost R., Berkowitz O., Wirtz M., Hopkins L., Hawkesford M.J., Hell R.;
RT RT "Genomic and functional characterization of the oas gene family
RT RT encoding O-acetylserine (thiol) lyases, enzymes catalyzing the final
RL RL step in cysteine biosynthesis in Arabidopsis thaliana.";
RM RM Gene 253:237-247(2000).
RN RN [5]
RP RP SEQUENCE FROM N.A.
RC RC STRAIN=cv. Columbia;
RX RX MEDLINE=20083487; PubMed=10617197;
RA RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
RA RA Tallon L.J., Gill J.B., Adams M.D., Carrera A.J., Creasy T.H.,
RA RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA RA Venter J.C.;
RT RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT RT thaliana.";
RL RL Nature 402:761-768(1999).
RM RM -!- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +
CC CC acetate.
CC CC -!- COFACTOR: Pyridoxal phosphate.
CC CC -!- PATHWAY: Cysteine biosynthesis.
CC CC -!- SUBUNIT: Homodimer.
CC CC -!- SUBCELLULAR LOCATION: Chloroplast stroma; cytoplasm.
CC CC -!- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-
CC CC synthase family.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
DR DR EMBL; X80377; CAA56594.2; -
DR DR EMBL; X81698; CAA57344.1; -
DR DR EMBL; AJ271728; CAB71292.1; -
DR DR EMBL; AC002333; AAB64031.1; -
DR DR PIR; A84870; A84870.
DR DR PIR; S48695; S48695.
DR DR HSSP; P12674; LOAS.
DR DR SWISS-2DPAGE; P47999; ARATH.
DR DR InterPro; IPR001926; B6_enzyme_beta.
DR DR InterPro; IPR001216; Cys_synthase_BS.
DR DR InterPro; IPR005859; Cys_synthK.
DR DR InterPro; IPR005856; Cys_synthKM.
DR DR Pfam; PF00291; PALP; 1.
DR DR TIGRfams; TIGR01139; cySK; 1.
DR DR TIGRfams; TIGR01136; cySKM; 1.

```

[illegible]



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X64874; CAA46086.1; -.  
DR PIR; A43407; A43407.  
DR HSSP; p12674; 10AS.  
DR InterPro; IPR001926; B6\_enzyme\_beta.  
DR InterPro; IPR001216; Cys\_synthase\_BS.  
DR InterPro; IPR005859; Cys\_synthase.  
DR InterPro; IPR005856; Cys\_synthase.  
DR Pfam; PF00291; PALP; 1.  
DR TIGRFAMs; TIGR01139; cysK; 1.  
DR TIGRFAMs; TIGR01136; cysK; 1.  
DR PROSITE; PS00901; CYS\_SYNTHASE; 1.  
KW Transferase; Cysteine biosynthesis; Pyridoxal phosphate; Chloroplast;  
KW Transit peptide.  
FT TRANSIT 1 50 CHLOROPLAST (BY SIMILARITY).  
FT CHAIN 51 374 CYSTEINE SYNTHASE.  
FT BINDING 108 108 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
SQ SEQUENCE 374 AA; 39978 MW; 36A6D0CCEA23F1 CRC64;

Query Match 77.0%; Score 1249; DB 1; Length 374;  
Best Local Similarity 75.2%; Pred. No. 2.1e-78;  
Matches 240; Conservative 38; Mismatches 31; Indels 10; Gaps 1;

QY 3 VERSGIADVTTELIGKTPVLYINKLADGCVARVAAKLEIMEPCSSVKDRIGYSMTADAEE 62  
Db 62 IEGINIAEDVTQLIGNTPMVYINTIVKGCVANIAAKLEIMEPCSSVKDRIGYSMTADAEE 121  
QY 63 KGLITPGKSVLIBPTSGNTGIGLAFAAAGYKLIITMPASMSLERRIILLAFGARIVLT 122  
Db 122 KGLISPGKTVLVEPTSGNTGIGLAFAASRGYKLIITMPASMSLERRVILKAFGARIVLT 181  
QY 123 DPAKMKGAVOKABEILAKTPNAYILQGFENPANPKVYETTTGPEIWKSGSDGKIDAFVSG 182  
Db 182 DPAKMKGAVSKABEILANTPDAYILQGFDPNPNPKIHYETTTGPEIWEETKGIKIDILVAG 241  
QY 183 IGTGGTITGAGKYLKEONPNIKLIGVEPVESPVLSGGKPGPHKIQIGAGFIPGVLENTL 242  
Db 242 IGTGGTISGTGRLKEKNPNIKLIGVEPTBSNVLSGGKPGPHKIQIGAGFIPGVLENTL 291  
QY 243 LDEVVQISSDEALETAKLALKEGLFVGISSGAAAAAFOIAKRPENAGKLIYAVFPSPFG 302  
Db 292 MDEVTEISSDEAVETAKQALQEGLLVGISSGAAALAAIQVAKRPENAGKLIYAVFPSPFG 351  
QY 303 ERYLSSVLFESVRREAESM 321  
Db 352 ERYLSSILFQSIREECEKM 370

Search completed: May 18, 2004, 13:07:34  
Job time : 18 secs



Db 1 MAVERSGIAKDVTELLIGKTPLVYLNKLADGCVARVAAKLELMEPCSSVYKDRIGYSMIADA 60  
 QY 61 EEKGLITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILLAFGAEV 120  
 Db 61 EEKGLITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILLAFGAEV 120  
 QY 121 LTDPKAGMKGAVQKAEIILAKTPNAVYILQGFENPANPKVHYETTGPEIWKSGSDKIDAFV 180  
 Db 121 LTDPKAGMKGAVQKAEIILAKTPNAVYILQGFENPANPKVHYETTGPEIWKSGSDKIDAFV 180  
 QY 181 SGIGTGITGAGKYLKEQNPNKILIGVEPVESPVLSGGKPGPHKIQIGAGFIPGVLEV 240  
 Db 181 SGIGTGITGAGKYLKEQNPNKILIGVEPVESPVLSGGKPGPHKIQIGAGFIPGVLEV 240  
 QY 241 NLDEVVOISSDEALETAKLLALKEGLFVGISGGAATAAFQIAKRPENAGKLIIVAFPS 300  
 Db 241 NLDEVVOISSDEALETAKLLALKEGLFVGISGGAATAAFQIAKRPENAGKLIIVAFPS 300  
 QY 301 FGERYLSVLFESVRRRAESMTFEP 325  
 Db 301 FGERYLSVLFESVRRRAESMTFEP 325

RESULT 2

ID Q9FS27 PRELIMINARY; PRT; 325 AA.  
 AC Q9FS27;  
 DT 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
 DE Cytosolic cysteine synthase.  
 GN PCS-1.  
 OS Solanum tuberosum (potato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC lamids; Solanales; Solanaceae; Solanum.  
 OC NCBI\_TaxID=4113;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Dansyaku;  
 RA Maruyama A., Ishizawa K.;  
 RT "Cytosolic Cysteine Synthase from Potato Sprout."  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB029511; BAB20861.1; -  
 DR HSSP; P12674; 1OAS.  
 DR GO; GO:0004124; F:cysteine synthase activity; IEA.  
 DR GO; GO:0016829; F:lyase activity; IEA.  
 DR GO; GO:0006520; P:amino acid metabolism; IEA.  
 DR GO; GO:0006535; P:cysteine biosynthesis from serine; IEA.  
 DR InterPro; IPR001926; B6\_enzyme\_beta.  
 DR InterPro; IPR005859; Cys\_synthK.  
 DR InterPro; IPR005856; Cys\_synthK.  
 DR Pfam; PF00291; PALP; 1.  
 DR TIGRFAMS; TIGR01139; cysK; 1.  
 DR TIGRFAMS; TIGR01136; cysK; 1.  
 SQ SEQUENCE 325 AA; 34308 MW; 9B0D526562F6C785 CRC64;

Query Match 89.0%; Score 1445; DB 10; Length 325;  
 Best Local Similarity 85.8%; Pred. No. 2e-93;  
 Matches 279; Conservative 26; Mismatches 20; Indels 0; Gaps 0;

QY 1 MAVERSGIAKDVTELLIGKTPLVYLNKLADGCVARVAAKLELMEPCSSVYKDRIGYSMIADA 60  
 Db 1 MAVERSGIAKDVTELLIGKTPLVYLNKLADGCVARVAAKLELMEPCSSVYKDRIGYSMIADA 60  
 QY 61 EEKGLITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILLAFGAEV 120  
 Db 61 EEKGLITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILLAFGAEV 120  
 QY 121 LTDPKAGMKGAVQKAEIILAKTPNAVYILQGFENPANPKVHYETTGPEIWKSGSDKIDAFV 180  
 Db 121 LTDPKAGMKGAVQKAEIILAKTPNAVYILQGFENPANPKVHYETTGPEIWKSGSDKIDAFV 180  
 QY 121 LTDPKAGMKGAVQKAEIILAKTPNAVYILQGFENPANPKVHYETTGPEIWKSGSDKIDAFV 180  
 Db 121 LTDPKAGMKGAVQKAEIILAKTPNAVYILQGFENPANPKVHYETTGPEIWKSGSDKIDAFV 180

QY 181 SGIGTGITGAGKYLKEQNPNKILIGVEPVESPVLSGGKPGPHKIQIGAGFIPGVLEV 240  
 Db 181 SGIGTGITGAGKYLKEQNPNKILIGVEPVESPVLSGGKPGPHKIQIGAGFIPGVLEV 240  
 QY 241 NLDEVVOISSDEALETAKLLALKEGLFVGISGGAATAAFQIAKRPENAGKLIIVAFPS 300  
 Db 241 NLDEVVOISSDEALETAKLLALKEGLFVGISGGAATAAFQIAKRPENAGKLIIVAFPS 300  
 QY 301 FGERYLSVLFESVRRRAESMTFEP 325  
 Db 301 FGERYLSVLFESVRRRAESMTFEP 325

RESULT 3

ID Q9MAZ2 PRELIMINARY; PRT; 325 AA.  
 AC Q9MAZ2;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
 DE Cysteine synthase.  
 GN BSAS1.  
 OS Allium tuberosum (Garlic chives).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; liliopsida; Asparagales; Alliaceae;  
 OC Allium.  
 OC NCBI\_TaxID=4683;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20534797; PubMed=11080593;  
 RA Urano Y., Manabe T., Noji M., Saito K.;  
 RT "Molecular cloning and functional characterization of cDNAs encoding  
 RT cysteine synthase and serine acetyltransferase that may be responsible  
 RT for high cellular cysteine content in Allium tuberosum."  
 RL Gene 257:269-277(2000).  
 DR EMBL; AB040503; BAA93051.1; -  
 DR HSSP; P12674; 1OAS.  
 DR GO; GO:0004124; F:cysteine synthase activity; IEA.  
 DR GO; GO:0016829; F:lyase activity; IEA.  
 DR GO; GO:0006520; P:amino acid metabolism; IEA.  
 DR GO; GO:0006535; P:cysteine biosynthesis from serine; IEA.  
 DR InterPro; IPR001926; B6\_enzyme\_beta.  
 DR InterPro; IPR001216; Cys\_synthase\_B5.  
 DR InterPro; IPR005859; Cys\_synthK.  
 DR InterPro; IPR005856; Cys\_synthK.  
 DR Pfam; PF00291; PALP; 1.  
 DR TIGRFAMS; TIGR01139; cysK; 1.  
 DR TIGRFAMS; TIGR01136; cysK; 1.  
 DR PROSITE; PS00901; CYS\_SYNTHASE; 1.  
 SQ SEQUENCE 325 AA; 34337 MW; C86BD91205DF358F CRC64;

Query Match 84.6%; Score 1373; DB 10; Length 325;  
 Best Local Similarity 82.4%; Pred. No. 2.3e-88;  
 Matches 262; Conservative 32; Mismatches 24; Indels 0; Gaps 0;

QY 8 IAKDVTELLIGKTPLVYLNKLADGCVARVAAKLELMEPCSSVYKDRIGYSMIADAERGLIT 67  
 Db 8 IAKDVTELLIGKTPLVYLNKLADGCVARVAAKLELMEPCSSVYKDRIGYSMIADAERGLIT 67  
 QY 68 PGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILLAFGAEVLTDPKAG 127  
 Db 68 PGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILLAFGAEVLTDPKAG 127  
 QY 128 MKGAVQKAEIILAKTPNAVYILQGFENPANPKVHYETTGPEIWKSGSDKIDAFVSGITGG 187  
 Db 128 MKGAVQKAEIILAKTPNAVYILQGFENPANPKVHYETTGPEIWKSGSDKIDAFVSGITGG 187  
 QY 188 TITGAGKYLKEQNPNKILIGVEPVESPVLSGGKPGPHKIQIGAGFIPGVLEVNLDEVV 247  
 Db 188 TITGAGKYLKEQNPNKILIGVEPVESPVLSGGKPGPHKIQIGAGFIPGVLEVNLDEVV 247  
 QY 248 OISSDEALETAKLLALKEGLFVGISGGAATAAFQIAKRPENAGKLIIVAFPSFGERYLS 307  
 Db 248 OISSDEALETAKLLALKEGLFVGISGGAATAAFQIAKRPENAGKLIIVAFPSFGERYLS 307



Db 248 QVSSSEAIEMAKOLALQEGILVIGISSGAAAIARVAQRPENEKLIIVVFPFGERYLS 307  
QY 308 SYLFESVRRRBSMTFEP 325  
Db 308 TVLFOSIKKEABSMVFEP 325

RESULT 4

Q94AS7 PRELIMINARY; PRT; 321 AA.  
AC Q94AS7;  
DT 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Putative cytosolic O-acetylserine(thiol)lyase.  
GN AT4G14880.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Yamada K., Liu S.X., Pham P.K., Banh J., Dale J.M., Goldsmith A.D.,  
RA Jiang P.X., Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M.,  
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,  
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,  
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,  
RA Shinn P., Southwick A., Tracy S.E., Shinzaki K., Davis R.W.,  
RA Becker J.R., Theologis A.;  
RT "Full length cDNA of gene AT4G14880 (GI:7268234).";  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY045825; AAK76499.1; -  
DR GO; GO:0004124; P:cysteine synthase activity; IEA.  
DR GO; GO:0016829; F:lyase activity; IEA.  
DR GO; GO:0006520; P:amino acid metabolism; IEA.  
DR GO; GO:0006535; P:cysteine biosynthesis from serine; IEA.  
DR InterPro; IPR001926; B6\_enzyme\_beta.  
DR InterPro; IPR001216; Cys\_synthase\_BS.  
DR InterPro; IPR005859; Cys\_synthK.  
DR InterPro; IPR005856; Cys\_synthKM.  
DR Pfam; PF00291; PALP; 1.  
DR TIGRFAMS; TIGR01139; cysK; 1.  
DR TIGRFAMS; TIGR01136; cysKM; 1.  
DR PROSITE; PS00901; CYS\_SYNTHASE; 1.  
KM Lyase.  
SQ SEQUENCE 321 AA; 33642 MW; 56531548B939B32B CRC64;

Query Match 84.2%; Score 1366.5; DB 10; Length 321;  
Best Local Similarity 83.1%; Pred. No. 6.4e-88;  
Matches 265; Conservative 30; Mismatches 23; Indels 1; Gaps 1;

QY 6 SGIAKDVTELIQKTPLVYLKLAAGCVAARVAAKLILMEPCSSVKDRIGYSMIADAEKGL 65  
Db 3 SRIKADYTELIGNTPLV-LNNVABGCVGRVAAKLEMEPCSSVKDRIGFSMISDAERKGL 61  
QY 66 ITPGKSVLIEPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILLAFGAELVLTDP 125  
Db 62 IKPGESVLIEPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILLAFGAELVLTDP 121  
QY 126 KGMKGAQVQKAEIILAKTPNAYILQGFENPANPKVHYETTGPEIWKSGDGKIDAFVSGIGT 185  
Db 122 KGMKGAIAKAEIILAKTPNGVYMLQGFENPANPKIHYETTGPEIWKSGDGKIDGFSVSGIGT 181  
QY 186 GGTITGAGKYLLKQNPNIKLIGVBPVESPVLSGKPGPHKIQGTIGAGFIPGVLEVLNLLDE 245  
Db 182 GGTITGAGKYLLKQNPANVKLYGVEPVESALISGKPGPHKIQGTIGAGFIPSVLANVLDIDE 241  
QY 246 VVQISSDEAIEIAKLALKEGLFVGISSGAAAAAFOIAKRPENAGKLIIVAVPSPGERY 305  
Db 242 VVQVSSDESIDMARQLALKEGLLVGISSGAAAAAIFKLAQRPENAGKLIIVAVPSPGERY 301

QY 306 LSSVLFESVRRRBSMTFEP 324  
Db 302 LSTVLFDAIRKEAAMTFEP 320

RESULT 5

Q9FS26 PRELIMINARY; PRT; 386 AA.  
AC Q9FS26;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Plastidic cysteine synthase 1.  
GN PCS-2-1.  
OS Solanum tuberosum (Potato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Lamiales; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4113;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Dansyaku; TISSUE=Sprout;  
RA Maruyama A., Ishizawa K.;  
RT "Plastidic Cysteine Synthase 1 (PCS-2-1) from Potato Sprout.";  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB029512; BAB20862.1; -  
DR HSP; P12674; 10AS.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0004124; P:cysteine synthase activity; IEA.  
DR GO; GO:0016829; F:lyase activity; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0005215; P:transporter activity; IEA.  
DR GO; GO:0006520; P:amino acid metabolism; IEA.  
DR GO; GO:0006535; P:cysteine biosynthesis from serine; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR001926; B6\_enzyme\_beta.  
DR InterPro; IPR001216; Cys\_synthase\_BS.  
DR InterPro; IPR005859; Cys\_synthK.  
DR InterPro; IPR005856; Cys\_synthKM.  
DR InterPro; IPR000531; TonE\_boxC.  
DR Pfam; PF00291; PALP; 1.  
DR TIGRFAMS; TIGR01139; cysK; 1.  
DR TIGRFAMS; TIGR01136; cysKM; 1.  
DR PROSITE; PS00901; CYS\_SYNTHASE; 1.  
DR PROSITE; PS00430; TONB\_DEPENDENT\_REC\_1; 1.  
SQ SEQUENCE 386 AA; 40519 MW; 8294B5B9D8640BC7 CRC64;

Query Match 81.1%; Score 1316; DB 10; Length 386;  
Best Local Similarity 78.4%; Pred. No. 2.9e-84;  
Matches 250; Conservative 37; Mismatches 32; Indels 0; Gaps 0;

QY 3 VERSGIAKDVTELIQKTPLVYLKLAAGCVAARVAAKLILMEPCSSVKDRIGYSMIADABE 62  
Db 64 IEGNLIAEDVTQLIGNTPMVYLTNTIAKGCYANIAAKLILMEPCSSVKDRIGFSMIVDABE 123  
QY 63 KGLITPGKSVLIEPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILLAFGAELVLT 122  
Db 124 KGLISPGKTVLIEPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILLAFGAELVLT 183  
QY 123 DPAKMGKGAQVQKAEIILAKTPNAYILQGFENPANPKVHYETTGPEIWKSGDGKIDAFVSG 182  
Db 184 DPAKMGKGAQVQKAEIILAKTPNAYILQGFENPANPKIHYETTGPEIWKSGDGKIDAFVSG 243  
QY 183 IGTGTTGAGKYLLKQNPNIKLIGVBPVESPVLSGKPGPHKIQGTIGAGFIPGVLEVLNLL 242  
Db 244 IGTGTTGAGKYLLKQNPNIKLIGVBPVESPVLSGKPGPHKIQGTIGAGFIPGVLEVLNLL 303  
QY 243 LDEVVQISSDEAIEIAKLALKEGLFVGISSGAAAAAFOIAKRPENAGKLIIVAVPSPGER 302  
Db 304 MDEVTEISSDEAIEIAKLALKEGLLVGISSGAAAAAFOIAKRPENAGKLIIVAVPSPGER 363  
QY 303 ERYLSSVLFESVRRRBSMTFEP 321

DB 364 ERYLSSILFQSIREFCEKM 382

## RESULT 6

09LJAO PRELIMINARY; PRT; 305 AA.

AC 09LJAO; 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Cysteine synthase, O-acetylserine(thiol) lyase.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI\_TaxID=3702;

[1]

RP SEQUENCE FROM N.A.

RC STRAIN=Columbia;

RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;

RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Columbia;

RX MEDLINE=20363099; PubMed=10907853;

RA Nakamura Y.;

RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.

RT Sequence features of the regions of 4,251,695 bp covered by ninety P1,

RT TAC and BAC clones.";

RL DNA Res. 7:217-221(2000).

DR EMBL; AP000731; BAB01461.1; -

DR HSSP; P12674; 10AS.

DR GO; GO:0004124; F:cysteine synthase activity; IEA.

DR GO; GO:0016829; F:lyase activity; IEA.

DR GO; GO:0006520; P:amino acid metabolism; IEA.

DR GO; GO:0006535; P:cysteine biosynthesis from serine; IEA.

DR InterPro; IPR001926; B6\_enzyme\_beta.

DR InterPro; IPR001216; Cys\_synthase\_BS.

DR InterPro; IPR005859; Cys\_synthk.

DR InterPro; IPR005856; Cys\_synthk.

DR Pfam; PF00291; PALP; 1.

DR TIGRFAMs; TIGR01139; cysK; 1.

DR TIGRFAMs; TIGR01136; cysKM; 1.

DR PROSITE; PS00901; CYS\_SYNTHASE; 1.

KW lyase.

SQ SEQUENCE 305 AA; 32338 MM; 130F434FC47382C1 CRC64;

Query Match 79.9%; Score 1296; DB 10; Length 305;

Best Local Similarity 78.2%; Pred. No. 5.3e-83;

Matches 254; Conservative 28; Mismatches 23; Indels 20; Gaps 1;

QY 1 MAVERSGIAKDVTELIGTPEVYLNKADGCVARVAKLELMEPCSSVKDRIGYSMIDA 60

DB 1 MASVAPKIAKDVTELIGTPEVYLNKADGCVARVAKLELMEPCSSVKDRIGYSMIDA 60

QY 61 EERGLITPKSVLIBPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILAFGAELV 120

DB 61 EAKGLIKPGESVLIBPTSGNTGIGLAFMAAARGYKLIITMPASMSIERRIILAFGAELI 120

QY 121 LTDPKAKMGAVQKAEIILAKTPNAYILQOFENPANPKVHYETGTPEIWKSGDKIDAFV 180

DB 121 LTDPKAKMGAVQKAEIILAKTPNAYILQOFENPANPKVHYETGTPEIWKSGDKIDAFV 180

QY 181 SGIGTGTITGAGKYLKQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPIGVLEV 240

DB 171 SGIGTGTITGAGKYLKQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPIGVLEV 240

QY 241 NLIDEVVOISSDEAIEITAKLALKEGLFVGISSGAAAAAFQIAKRPENAGKLIIVAFPS 300

DB 221 DLIDEVVOISSDEAIEITAKLALKEGLFVGISSGAAAAAFQIAKRPENAGKLIIVAFPS 300

QY 301 FGERYLSVLFESVRRAEASMTFEP 325

DB 281 FGERYSTVLFDARKEAETMTFEP 305

## RESULT 7

065747 PRELIMINARY; PRT; 266 AA.

AC 065747; 01-AUG-1998 (TREMBlrel. 07, Created)

DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE Cysteine synthase, O-acetyl-L-serine (Thiol)-lyase (EC 4.2.99.8)

DE (Fragment).

OS Cicer arietinum (Chickpea) (Garbanzo).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids I; Fabales; Fabaceae; Papilionoideae; Ciceraceae; Cicer.

OX NCBI\_TaxID=3827;

[1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Castellana; TISSUE=Etisolated epicotyl;

RA Dopico B., Estebean R., Labrador E.;

RT "cdna expressed in chickpea epicotyls.";

RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ006024; CAA06819.1; -

DR HSSP; P12674; 1PCJ.

DR GO; GO:0004124; F:cysteine synthase activity; IEA.

DR GO; GO:0016829; F:lyase activity; IEA.

DR GO; GO:0006520; P:amino acid metabolism; IEA.

DR GO; GO:0006535; P:cysteine biosynthesis from serine; IEA.

DR InterPro; IPR001926; B6\_enzyme\_beta.

DR InterPro; IPR005859; Cys\_synthk.

DR InterPro; IPR005856; Cys\_synthk.

DR Pfam; PF00291; PALP; 1.

DR TIGRFAMs; TIGR01139; cysK; 1.

DR TIGRFAMs; TIGR01136; cysKM; 1.

KW lyase.

FT NON TER 1

SQ SEQUENCE 266 AA; 27911 MM; 1P253E511A3596B CRC64;

Query Match 79.5%; Score 1290; DB 10; Length 266;

Best Local Similarity 94.7%; Pred. No. 1.2e-82;

Matches 252; Conservative 12; Mismatches 2; Indels 0; Gaps 0;

QY 60 ABEKGLITPKSVLIBPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILAFGAEL 119

DB 1 ABEKGLITPKSVLIBPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILAFGAEL 119

QY 120 VLTDPKAKMGAVQKAEIILAKTPNAYILQOFENPANPKVHYETGTPEIWKSGDKIDAF 179

DB 61 VLTDPKAKMGAVQKAEIILAKTPNAYILQOFENPANPKVHYETGTPEIWKSGDKIDAF 179

QY 180 VSGIGTGTITGAGKYLKQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPIGVLE 239

DB 121 VSGIGTGTITGAGKYLKQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPIGVLE 239

QY 240 VNLIDEVVOISSDEAIEITAKLALKEGLFVGISSGAAAAAFQIAKRPENAGKLIIVAFV 299

DB 181 VNLIDEVVOISSDEAIEITAKLALKEGLFVGISSGAAAAAFQIAKRPENAGKLIIVAFV 299

QY 300 SFGERYLSVLFESVRRAEASMTFEP 325

DB 241 SFGERYLSVLFESVRRAEASMTFEP 325

## RESULT 8

084SE4 PRELIMINARY; PRT; 409 AA.

AC 084SE4; 01-JUN-2003 (TREMBlrel. 24, Created)

DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Putative plastidic cysteine synthase 1.

GN P0020E09.20.

Oryza sativa (japonica cultivar-group).  
 OC Burkaria; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Eubryotidae; Oryzae; Oryza.  
 NCBI\_TaxID=39947;

[1]  
 SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RX MEDLINE=2237376; PubMed=12447438;  
 RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,  
 RA Wu J., Nishimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,  
 RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,  
 RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C., Ikeda M.,  
 RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,  
 RA Ikono M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,  
 RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,  
 RA Machida K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,  
 RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,  
 RA Namiki N., Negishi M., Ohba I., Ono N., Saji S., Sakai K., Shibata M.,  
 RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,  
 RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,  
 RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,  
 RA Yano M., Jiang J., Gojobori T.,  
 RT "The genome sequence and structure of rice chromosome 1,"  
 RL Nature 420:312-316 (2002).  
 DR EMBL; AP003228; BAC57641.1;  
 DR GO; GO:0004124; F:cysteine synthase activity; IEA.  
 DR GO; GO:0016829; F:lyase activity; IEA.  
 DR GO; GO:0006520; P:amino acid metabolism; IEA.  
 DR GO; GO:0006535; P:cysteine biosynthesis from serine; IEA.  
 DR InterPro; IPR001926; B6\_enzyme\_beta.  
 DR InterPro; IPR001216; Cys\_synthase\_BS.  
 DR InterPro; IPR005859; Cys\_synthK.  
 DR InterPro; IPR005856; Cys\_synthKM.  
 DR Pfam; PF00291; PALP; 1.  
 DR TIGRfams; TIGR01139; cysK; 1.  
 DR TIGRfams; TIGR01136; cysKM; 1.  
 DR PROSITE; PS00901; CYS SYNTHASE; 1.  
 DR PROSITE; PS00901; CYS SYNTHASE; 1.  
 SQ SEQUENCE 409 AA; 43585 MW; BSA04875B5F9B897 CRC64;

Query Match 78.6%; Score 1275.5; DB 10; Length 409;  
 Best Local Similarity 72.7%; Pred. No. 2.2e-81;  
 Matches 248; Conservative 38; Mismatches 34; Indels 21; Gaps 2;

QY 3 VERSGIADVTLEIGTPTLVYLNKLDGCVARVAAKLELMEPCSSVKDRIGYSMIADAE 62  
 DB 72 VEGLNIAADD---LIGKTPMYLVNINVKGVANVAAKLEIMEPCSSVKDRIGYSMIADAE 128  
 QY 63 KGLITPGR-----SVLISPTSGNTGIGLAEMAARGYKLIITMPASM 104  
 DB 129 KGLITPGRKLSPSDLPCMTYYFMLVAQSVLWEPTSGNTGIGLAFLAASRGYKLIITMPASM 188  
 QY 105 SLERRIILAFGAELVLTDPAGKMGAVQKAEIILAKTPNAYILQOENPANPKVHYETT 164  
 DB 189 SMERRVILKAFGAELVLTDAAGKMGAVQKAEIILAKTPNAYILQOENPANPKVHYETT 248  
 QY 165 GPRIWGSQDKIDAFVSGITGTTGAGKYLKQNPNIKLIGVEPVESFVLSGKPGPH 224  
 DB 249 GPRIWGSQDKIDAFVSGITGTTGAGKYLKQNPNIKLIGVEPVESFVLSGKPGPH 308  
 QY 225 KIIGIGAGFIPVLEVNLLDEVVOISSDAIETAKLILAKELFGVIGSSGAAAFQIA 284  
 DB 309 KIIGIGAGFIPVLEVNLLDEVVOISSDAIETAKLILAKELFGVIGSSGAAAFQIA 368  
 QY 285 KRPNAGKLIYAVFPSPFGERYLSVLFESVRRRAESMTFEP 325  
 DB 369 KRPNAGKLIYAVFPSPFGERYLSVLFESVRRRAESMTFEP 409

DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Cysteine synthase.  
 GN ATG59760.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Burkaria; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 NCBI\_TaxID=3702;

[1]  
 SEQUENCE FROM N.A.  
 RP Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,  
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,  
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,  
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,  
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,  
 RA Becker J., Theologis A., Davis R.W.,  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RA Tripp M., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,  
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,  
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,  
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,  
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,  
 RA Becker J., Theologis A., Davis R.W.,  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY099721; AAM20572.1;  
 DR EMBL; AY128885; AAM91285.1;  
 DR PIR; T47800; T47800.  
 DR GO; GO:0004124; F:cysteine synthase activity; IEA.  
 DR GO; GO:0016829; F:lyase activity; IEA.  
 DR GO; GO:0006520; P:amino acid metabolism; IEA.  
 DR GO; GO:0006535; P:cysteine biosynthesis from serine; IEA.  
 DR InterPro; IPR001926; B6\_enzyme\_beta.  
 DR InterPro; IPR001216; Cys\_synthase\_BS.  
 DR InterPro; IPR005859; Cys\_synthK.  
 DR InterPro; IPR005856; Cys\_synthKM.  
 DR Pfam; PF00291; PALP; 1.  
 DR TIGRfams; TIGR01139; cysK; 1.  
 DR TIGRfams; TIGR01136; cysKM; 1.  
 DR PROSITE; PS00901; CYS SYNTHASE; 1.  
 DR PROSITE; PS00901; CYS SYNTHASE; 1.  
 SQ SEQUENCE 430 AA; 45814 MW; 9DDEB1B8B314E143 CRC64;

Query Match 78.3%; Score 1271; DB 10; Length 430;  
 Best Local Similarity 73.9%; Pred. No. 4.8e-81;  
 Matches 241; Conservative 43; Mismatches 36; Indels 6; Gaps 1;

QY 2 AVERS-----GIADVTLEIGTPTLVYLNKLDGCVARVAAKLELMEPCSSVKDRIGYS 55  
 DB 101 AVKRETPDGLINADVNVSQILGKTPMYLVNINVKGVANVAAKLEIMEPCSSVKDRIGYS 160  
 QY 56 MIADAEKGLITPGRKLSPSDLPCMTYYFMLVAQSVLWEPTSGNTGIGLAFLAASRGYKLIITMPASM 115  
 DB 161 MVTDAEQKFTSPKSVLWEPTSGNTGIGLAFLAASRGYKLIITMPASM 220  
 QY 116 GAELVLTDPAGKMGAVQKAEIILAKTPNAYILQOENPANPKVHYETTGPRIWGSQDK 175  
 DB 221 GAELVLTDPAGKMGAVQKAEIILAKTPNAYILQOENPANPKVHYETTGPRIWGSQDK 280  
 QY 176 IDAFVSGITGTTGAGKYLKQNPNIKLIGVEPVESFVLSGKPGPHKIOGIGAFIP 235  
 DB 281 VDIFFVAGITGTTGAGKYLKQNPNIKLIGVEPVESFVLSGKPGPHKIOGIGAFIP 340  
 QY 236 GVLVNLDEVVOISSDAIETAKLILAKELFGVIGSSGAAAFQIAKRPNAGKLIY 295  
 DB 341 KNIQKIMDEVVOISSDAIETAKLILAKELFGVIGSSGAAAFQIAKRPNAGKLIY 400  
 QY 296 AVFPSPFGERYLSVLFESVRRRAESMTFEP 321  
 DB 401 VVFPSPFGERYLSVLFESVRRRAESMTFEP 426



RESULT 10  
ID 081523 PRELIMINARY; PRT; 352 AA.  
AC 081523;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Putative O-acetylserine(thiol)lyase precursor (EC 4.2.99.8).  
GN CRCYS-1A.  
OS Chlamydomonas reinhardtii.  
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
OC Chlamydomonadaceae; Chlamydomonas.  
OX NCBI\_TaxID=3055;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99421691; PubMed=10491132;  
RA Ravina C.G., Barroso C., Vega J.M., Gotor C.,  
RT "Cysteine biosynthesis in Chlamydomonas reinhardtii. Molecular cloning  
and regulation of O-acetylserine(thiol)lyase.";  
RE Bur. J. Biochem. 264:848-853(1999).  
DR EMBL; AF078693; AAC27794.1; -.  
DR PIR; T07962; T07962.  
DR HSSP; P12674; 1OAS.  
DR GO; GO:0004124; F:cysteine synthase activity; IEA.  
DR GO; GO:0016829; F:lyase activity; IEA.  
DR GO; GO:0006520; P:amino acid metabolism; IEA.  
DR GO; GO:0006535; P:cysteine biosynthesis from serine; IEA.  
DR InterPro; IPR001926; B6\_enzyme\_beta.  
DR InterPro; IPR001216; Cys\_synthase\_BS.  
DR InterPro; IPR005859; Cys\_synthK.  
DR InterPro; IPR005856; Cys\_synthKM.  
DR Pfam; PF00291; PALP; 1.  
DR TIGRFAMS; TIGR01139; cysK; 1.  
DR TIGRFAMS; TIGR01136; cysKM; 1.  
DR PROSITE; PS00901; CYS\_SYNTHASE; 1.  
KW lyase; transit peptide.  
FT TRANSIT 1 31 POTENTIAL.  
FT CHAIN 32 352 POTENTIAL.  
SQ SEQUENCE 352 AA; 36858 MW; 6A48C19231FEE79F CRC64;  
Query Match 78.3%; Score 1270; DB 10; Length 352;  
Best Local Similarity 74.4%; Pred. No. 4.3e-81;  
Matches 241; Conservative 41; Mismatches 42; Indels 0; Gaps 0;  
QY 2 AVERSGIAKDVTELGKTPLVYLNKLADGCYARVAAKLEIMPCSSVKDRIGYSMTADAE 61  
DB 27 AAVKKNIAITDVTELGKTPMVYLNKVAITGTHAKIAAKLEIMPCSSVKDRIGYSMTADAE 86  
QY 62 EKGLITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILAFGAEVLT 121  
DB 87 KEGLITPGKTVLVEPTSGNTGIGLAFIAAARGYKLIITMPASMSLERRIILAFGAEVLT 146  
QY 122 TDPAKMKGAVQAKAEILAKTPNAVILQOFENPANPKVHYETTGPEIWKSGDKIDAFVS 181  
DB 147 TDPAKMKGAVQAKAEILAKTPNAVILQOFENPANPKVHYETTGPEIWKSGDKIDAFVS 206  
QY 182 GIGTGTITGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGIGAGFITPGVLEVN 241  
DB 207 GVGTTGTITGTRFLKEKSDVQLVAVEPAESPVLSGKPGPHKIQIGIGAGFITPGVLEVN 266  
QY 242 LDEVVOISSDEALETAKLILKEGLFVGISGAAAAAFAQIAKRPENAGKLIIVAPPSFG 301  
DB 267 LISEVVOYSSDDAIDMARLILKEGLMVGISGAAVQAAIKVASRPNENEGKLIVVVLDPF 326  
QY 302 GERYSVLVFSVRRAESMTFEP 325  
DB 327 GERYSVLVFOQLRDEASKMTFEP 350

RESULT 12  
ID 043726 PRELIMINARY; PRT; 324 AA.  
AC 043726;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Plastidic cysteine synthase 2.  
GN PCS-2-2.  
OS Solanum tuberosum (Potato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Lamiales; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4113;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Dansyaku; TISSUE=Sprout;  
RA Maruyama A., Ishizawa K.;  
RT "Plastidic Cysteine Synthase (PCS-2-2) from Potato Sprout.";  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB029513; BAB20863.1; -.  
DR HSSP; P12674; 1OAS.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0004124; F:cysteine synthase activity; IEA.  
DR GO; GO:0016829; F:lyase activity; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0005215; P:transporter activity; IEA.  
DR GO; GO:0006520; P:amino acid metabolism; IEA.  
DR GO; GO:0006535; P:cysteine biosynthesis from serine; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR001926; B6\_enzyme\_beta.  
DR InterPro; IPR001216; Cys\_synthase\_BS.  
DR InterPro; IPR005859; Cys\_synthK.  
DR InterPro; IPR005856; Cys\_synthKM.  
DR InterPro; IPR00531; TonB\_boxC.  
DR Pfam; PF00291; PALP; 1.  
DR TIGRFAMS; TIGR01139; cysK; 1.  
DR TIGRFAMS; TIGR01136; cysKM; 1.  
DR PROSITE; PS00901; CYS\_SYNTHASE; 1.  
DR PROSITE; PS00430; TONB\_DEPENDENT\_REC\_1; 1.  
SQ SEQUENCE 376 AA; 39960 MW; 47FDDA6EE2C6A75C CRC64;  
Query Match 76.8%; Score 1246; DB 10; Length 376;  
Best Local Similarity 75.2%; Pred. No. 2.3e-79;  
Matches 240; Conservative 37; Mismatches 32; Indels 10; Gaps 1;  
QY 3 VERSGIAKDVTELGKTPLVYLNKLADGCYARVAAKLEIMPCSSVKDRIGYSMTADAE 62  
DB 64 IEGLNIAEDVTOGLNTPMVYLNKIAKGVANIAAKLEIMPCSSVKDRIGYSMTADAE 123  
QY 63 KGLITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILAFGAEVLT 122  
DB 124 KGLISPKTVLVEPTSGNTGIGLAFIAAARGYKLIITMPASMSLERRIILAFGAEVLT 183  
QY 123 DPAKMKGAVQAKAEILAKTPNAVILQOFENPANPKVHYETTGPEIWKSGDKIDAFVS 182  
DB 184 DPAKMKGAVQAKAEILAKTPNAVILQOFENPANPKVHYETTGPEIWKSGDKIDAFVS 243  
QY 183 IGTGTTITGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGIGAGFITPGVLEVN 242  
DB 244 IGTGTTITGTRFLKEKSDVQLVAVEPAESPVLSGKPGPHKIQIGIGAGFITPGVLEVN 293  
QY 243 LDEVVOISSDEALETAKLILKEGLFVGISGAAAAAFAQIAKRPENAGKLIIVAPPSFG 302  
DB 294 MDEVIBISSDEALETAKLILKEGLFVGISGAAAAAFAQIAKRPENAGKLIIVAPPSFG 353  
QY 303 ERYLSSVLVFSVRRAESMTFEP 321  
DB 354 ERYLSSVLVFOQLRDEASKMTFEP 372

01-NOV-1996 (TRENBLREL. 01, Last sequence update)  
 01-JUN-2003 (TRENBLREL. 24, Last annotation update)  
 Cysteine synthase (EC 4.2.99.8).  
 GN Cyt ACS 1.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia; TISSUE=flower;  
 RX MEDLINE=95334513; PubMed=7610184;  
 RA Hesse H., Altman T.;  
 RT "Molecular cloning of a cysteine synthase cDNA from Arabidopsis  
 thaliana.";  
 RL Plant Physiol. 108:851-852(1995).  
 DR EMBL; X81697; CAA57343.1; -.  
 DR PIR; S49586; S49586.  
 DR HSSP; P12674; 10AS.  
 DR GO; GO:0004124; F:cysteine synthase activity; IEA.  
 DR GO; GO:0016829; F:lyase activity; IEA.  
 DR GO; GO:0006520; P:amino acid metabolism; IEA.  
 DR GO; GO:0006535; P:cysteine biosynthesis from serine; IEA.  
 DR InterPro; IPR001926; B6\_enzyme\_beta.  
 DR InterPro; IPR005859; Cys\_synth.  
 DR InterPro; IPR005856; Cys\_synth.  
 DR Pfam; PF00291; PALP; 1.  
 DR TIGRFAMs; TIGR01139; cysK; 1.  
 DR TIGRFAMs; TIGR01136; cysK; 1.  
 KW lyase.  
 SQ SEQUENCE 324 AA; 34605 MW; 6F43A718C1B3E1BD CRC64;  
 Query Match 75.7%; Score 1229; DB 10; Length 324;  
 Best Local Similarity 75.9%; Pred. No. 2.9e-78;  
 Matches 246; Conservative 34; Mismatches 38; Indels 6; Gaps 4;  
 QY 6 SGIAKDVTELLGKTPPLVYLKLA-DGCVARVAAKLEMEPCS-SVKDRIGYSMTDAEE 62  
 DB 3 SRIAKDVTELLGKTPPLVYLNNVSLKGCVRVAAKLEMEPCS-SVSKDRIGYSMTDAEK 62  
 QY 63 KGLITPGKSVLEPTSGNTG-IGLAFMAAARGYKLIITMPASMSLERRIILAFGAELVL 121  
 DB 63 KGLIKQGESVLEPTSGNTGELLAFMAACKGYKLIITMPASMSLERRIILAFGVEVL 122  
 QY 122 TDPAKMGKAVQKAEIILAKTPNAVILQGFENPANPKVHYETGPEIWKSGDKIDAFVS 181  
 DB 123 TDPAKMGKSYKAEIILAKTPNGVYLQGFENPANPEIHYETGPEIWKSGAKIDGFS 182  
 QY 182 GIGTGTTGAGKYLKEQNPNKILGVEPVESPVLSGKPGPHKIOGIGAGFIPGVLEVN 241  
 DB 183 GIGTGTTGAGKYLKEQNPNFKLYGVEPVESAILSGKPGPTKIOGIGAGFIPSVLNV 242  
 QY 242 LLDENVQISSDEALETAKLLALKEGLFVGISGAAAFQIAKRPENAGKLIVAVPFSF 301  
 DB 243 LIDENVQSSDESIDMARQLALKEGLVGISG--AAAFKLAQRPENAGKLIVAVPFSF 300  
 QY 302 GERYSVLFESVRRLEASMTFEP 325  
 DB 301 GRRYLSVLFADKERSRNMPSRP 324  
 RESULT 13  
 Q8W313 PRELIMINARY; PRT; 282 AA.  
 AC Q8W313;  
 DT 01-MAR-2002 (TRENBLREL. 20, Created)  
 DT 01-MAR-2002 (TRENBLREL. 20, Last sequence update)  
 DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)  
 DE Cysteine synthase, 5'-partial (Fragment).  
 GN OSJNB0069E14.1.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzae; Oryza.  
 OC NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Moffat K.S., Hill J.N.,  
 RA Gansberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,  
 RA Taitlin T., Riggs F., Hsiao J., Zismann V., Blunt S., Pai G.,  
 RA Vanaken S.E., Utterback T.R., Feldblyum T.V., Kalb E., Quackenbush J.,  
 RA Salzberg S.L., White O., Fraser C.M.;  
 RT "Oryza sativa chromosome 3 BAC OSJNB0069E14 genomic sequence.";  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC091811; AAL58961.1; -.  
 DR Gramene; Q8W313; -.  
 DR GO; GO:0004124; F:cysteine synthase activity; IEA.  
 DR GO; GO:0016829; F:lyase activity; IEA.  
 DR GO; GO:0006520; P:amino acid metabolism; IEA.  
 DR GO; GO:0006535; P:cysteine biosynthesis from serine; IEA.  
 DR InterPro; IPR001926; B6\_enzyme\_beta.  
 DR InterPro; IPR005859; Cys\_synth.  
 DR InterPro; IPR005856; Cys\_synth.  
 DR Pfam; PF00291; PALP; 1.  
 DR TIGRFAMs; TIGR01139; cysK; 1.  
 DR TIGRFAMs; TIGR01136; cysK; 1.  
 FT NON TER 1  
 SQ SEQUENCE 282 AA; 29892 MW; 3503D2F3CA80EFB5 CRC64;  
 Query Match 75.6%; Score 1227; DB 10; Length 282;  
 Best Local Similarity 82.7%; Pred. No. 3.3e-78;  
 Matches 235; Conservative 24; Mismatches 23; Indels 2; Gaps 1;  
 QY 42 MEPCSSVKDRIGYSMTADABEKLIITPGKSVLEPTSGNTGIGLAFMAAARGYKLIITMP 101  
 DB 1 MEPCSSVKDRIGYSMTADABEKLIITPGKSVLEPTSGNTGIGLAFMAAARGYKLIITMP 60  
 QY 102 ASMSLEERRIILAFGAELVLTDPKMGKAVQKAEIILAKTPNAVILQGFENPANPKVHY 161  
 DB 61 ASMSMERRIILKAFGAELVLTDPKMGKAVQKAEIILAKTPNAVILQGFENPANPKVHY 120  
 QY 162 ETTGPEIWKSGDKIDAFVSGIGTGTTGAGKYLKEQNPNKILGVEPVESPVLSGKPG 221  
 DB 121 ETTGPEIWKATAGKVDILVSGIGTGTTGAGKYLKEQNPNKILGVEPVESPVLSGKPG 180  
 QY 222 GPHKIOGIGAGFIPGVLEVNLLDEVVOISSDEALETAKLLALKEGLFVGISGAAAF 281  
 DB 181 GPHKIOGIGAGFIPGVLDVNLDEVVOSSDEAISMALKEGLLVGISGAAAF 240  
 QY 282 QIAKRPENAGKLIVAVPFSFGERYSVLFESVRRLEASMTFEP 325  
 DB 241 RVAQRPENAGKLIVAVPFSFGERYSVLFESIKREAENNVFEP 282  
 RESULT 14  
 Q9FSF5 PRELIMINARY; PRT; 324 AA.  
 AC Q9FSF5;  
 DT 01-MAR-2001 (TRENBLREL. 16, Created)  
 DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)  
 DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)  
 DE Cysteine synthase.  
 GN OS1.  
 OS Nicotiana tabacum (Common tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC lamids; Solanales; Solanaceae; Nicotiana.  
 OC NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Shen W.H.;  
 RT "Nicotiana tabacum cDNA encoding cysteine synthase (O-acetylserine  
 sulphydrylase).";  
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ299249; CAC12819.1; -  
 DR HSSP; P12674; 10AS.  
 DR GO; GO:0004124; F:cysteine synthase activity; IEA.  
 DR GO; GO:0016829; F:lyase activity; IEA.  
 DR GO; GO:0006520; P:amino acid metabolism; IEA.  
 DR GO; GO:0006535; P:cysteine biosynthesis from serine; IEA.  
 DR InterPro; IPR01926; B6\_enzyme\_beta.  
 DR InterPro; IPR01216; Cys\_synthase\_BS.  
 DR InterPro; IPR05859; Cys\_synthK.  
 DR InterPro; IPR05856; Cys\_synthKM.  
 DR Pfam; PF00291; PALP; 1.  
 DR TIGRfams; TIGR01139; cysK; 1.  
 DR TIGRfams; TIGR01136; cysKM; 1.  
 DR PROSITE; PS00901; CYS\_SYNTHASE; 1.  
 DR SEQUENCE 324 AA; 34958 MW; 328C3789D2983BA4 CRC64;

Query Match 75.2%; Score 1220; DB 10; Length 324;  
 Best Local Similarity 71.7%; Pred. No. 1.2e-77;  
 Matches 230; Conservative 45; Mismatches 46; Indels 0; Gaps 0;

QY 4 ERSGLAKDVTLLIGTPTLVYLNKLDGCVARVAAKLEMEPCSSVKDRIGYSMIDAEK 63  
 DB 3 EKSDIAKDVTELLIGTPTLVYLNKLDGCVARVAAKLEMEPCSSVKDRIGYSMIDAEK 62  
 QY 64 GLITPGKSVLIEPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILAFGAELVLT 123  
 DB 63 GLITPGKSVLIEPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILAFGAELVLT 122  
 QY 124 PAKGKAVOKABEILAKTPNAYILQOFENPANPKVHYETTPPEIWKSGDKIDAFVSGI 183  
 DB 123 PAKGIDGVLQKABEILAKTPNAYILQOFENPANPKVHYETTPPEIWKSGDKIDAFVSGI 182  
 QY 184 GTGGTITGAGKYLKEQNPENIKLIGVEPVESPVLSGKPGPHKIQIGAGIPLGVLEVNLL 243  
 DB 183 GTGGTITGAGKYLKEQNPENIKLIGVEPVESPVLSGKPGPHKIQIGAGIPLGVLEVNLL 242  
 QY 244 DEVVOISSDEAIEITAKLALKEGLFVGISGAAAGAAAFQIAKRPENAGKLIIVAFPSFGE 303  
 DB 243 DEVVOISSDEAIEITAKLALKEGLFVGISGAAAGAAAFQIAKRPENAGKLIIVAFPSFGE 302  
 QY 304 RYLSVLFESVRRBAESMTPE 324  
 DB 303 RYLSVLFESVRRBAESMTPE 323

RESULT 15

Q9SXS7 PRELIMINARY; PRT; 323 AA.  
 AC Q9SXS7;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
 DE Cysteine synthase (EC 4.2.99.8) (AT5g28020/f15f15\_90).  
 GN ATCYS2.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA Yamaguchi Y., Nakamura T., Kusano T., Sano H.;  
 RT "Cysteine synthases from Arabidopsis thaliana: Differential enzymatic  
 activities and distinct subcellular localization.";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,  
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,  
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,  
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,  
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,

RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,  
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,  
 RA Beker J.R.;  
 RT "Arabidopsis cDNA clones.";  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Kim C.J., Chen H., Cheuk R., Shin P., Banh J., Bowser L.,  
 RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,  
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,  
 RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,  
 RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
 RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,  
 RA Shinozaki K., Davis R.W., Theologis A., Beker J.R.;  
 RT "Arabidopsis ORF clones.";  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB024283; BAA78561.1; -  
 DR EMBL; AF424598; AAL11592.1; -  
 DR EMBL; AY124831; AAM70540.1; -  
 DR HSSP; P12674; 10AS.  
 DR GO; GO:0004124; F:cysteine synthase activity; IEA.  
 DR GO; GO:0016829; F:lyase activity; IEA.  
 DR GO; GO:0006520; P:amino acid metabolism; IEA.  
 DR GO; GO:0006535; P:cysteine biosynthesis from serine; IEA.  
 DR InterPro; IPR01926; B6\_enzyme\_beta.  
 DR InterPro; IPR01216; Cys\_synthase\_BS.  
 DR InterPro; IPR05859; Cys\_synthK.  
 DR InterPro; IPR05856; Cys\_synthKM.  
 DR Pfam; PF00291; PALP; 1.  
 DR TIGRfams; TIGR01139; cysK; 1.  
 DR TIGRfams; TIGR01136; cysKM; 1.  
 DR PROSITE; PS00901; CYS\_SYNTHASE; 1.  
 DR KW Lyase.  
 DR SEQUENCE 323 AA; 34318 MW; 877C8677BEA777F4 CRC64;

Query Match 73.4%; Score 1192; DB 10; Length 323;  
 Best Local Similarity 70.7%; Pred. No. 1.1e-75;  
 Matches 227; Conservative 45; Mismatches 49; Indels 0; Gaps 0;

QY 4 ERSGLAKDVTLLIGTPTLVYLNKLDGCVARVAAKLEMEPCSSVKDRIGYSMIDAEK 63  
 DB 3 DRCLINQDITELLIGTPTLVYLNKLDGCVARVAAKLEMEPCSSVKDRIGYSMIDAEK 62  
 QY 64 GLITPGKSVLIEPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILAFGAELVLT 123  
 DB 63 GLITPGKSVLIEPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILAFGAELVLT 122  
 QY 124 PAKGKAVOKABEILAKTPNAYILQOFENPANPKVHYETTPPEIWKSGDKIDAFVSGI 183  
 DB 123 ORIGLKNLEKTEALISKTGGYIFQOFENPANPEIHYRTTGPPEIWRDSAGKVDILVAGV 182  
 QY 184 GTGGTITGAGKYLKEQNPENIKLIGVEPVESPVLSGKPGPHKIQIGAGIPLGVLEVNLL 243  
 DB 183 GTGGTITGAGKYLKEQNPENIKLIGVEPVESPVLSGKPGPHKIQIGAGIPLGVLEVNLL 242  
 QY 244 DEVVOISSDEAIEITAKLALKEGLFVGISGAAAGAAAFQIAKRPENAGKLIIVAFPSFGE 303  
 DB 243 DEIIQVAGERAIEITAKLALKEGLFVGISGAAAGAAAFQIAKRPENAGKLIIVAFPSFGE 302  
 QY 304 RYLSVLFESVRRBAESMTPE 324  
 DB 303 RYLSVLFESVRRBAESMTPE 323

Search completed: May 18, 2004, 13:08:33  
 Job time : 47 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: May 18, 2004, 13:05:59 ; Search time 23 Seconds  
(without alignments)  
729.497 Million cell updates/sec

Title: US-09-931-457A-31  
Perfect score: 1623  
Sequence: 1 MAVERSGIAKDVTELGKTP.....LSSVLFESVVRRAESMTFEP 325

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1179	72.6	324	US-09-905-290A-4	Sequence 4, Appli
2	882	54.3	335	US-09-543-681A-7733	Sequence 7733, Ap
3	865	53.3	335	US-09-328-352-5450	Sequence 5450, Ap
4	856	52.7	748	US-09-252-991A-31491	Sequence 31491, A
5	853.5	52.6	326	US-09-489-039A-9166	Sequence 9166, Ap
6	813	50.1	308	US-09-107-532A-3925	Sequence 3925, Ap
7	801	49.4	319	US-09-134-001C-3330	Sequence 3330, Ap
8	761	46.9	309	US-09-724-623-72	Sequence 72, Appl
9	650.5	40.1	303	US-09-543-681A-7720	Sequence 7720, Ap
10	628.5	38.7	306	US-09-489-039A-9175	Sequence 9175, Ap
11	598.5	36.9	382	US-09-252-991A-29566	Sequence 29566, A
12	594	36.6	307	US-09-328-352-8179	Sequence 8179, Ap
13	593	36.5	310	US-09-540-236-3342	Sequence 3342, Ap
14	584	36.0	378	US-09-789-300A-4	Sequence 4, Appli
15	552.5	34.0	308	US-09-107-532A-6629	Sequence 6629, Ap
16	550	33.9	551	US-08-120-960-2	Sequence 2, Appli
17	550	33.9	551	US-09-347-878-9	Sequence 9, Appli
18	541.5	33.4	1252	US-10-012-762-20	Sequence 20, Appl
19	541.5	33.4	1252	US-09-704-036B-20	Sequence 20, Appl
20	526	32.4	279	US-09-134-000C-5401	Sequence 5401, Ap
21	485	29.9	458	US-09-489-039A-9184	Sequence 9184, Ap
22	481	29.6	478	US-09-252-991A-16935	Sequence 16935, A
23	474	29.2	312	US-09-134-001C-3920	Sequence 3920, Ap
24	461	28.4	196	US-09-134-000C-3526	Sequence 3526, Ap
25	451.5	27.8	311	US-09-252-991A-25027	Sequence 25027, A
26	408	25.1	179	US-09-621-976-4379	Sequence 4379, Ap
27	369.5	22.8	343	US-09-543-681A-5853	Sequence 5853, Ap

28	290.5	17.9	137	4	US-09-134-000C-3525	Sequence 3525, Ap
29	233.5	14.4	359	4	US-09-489-039A-13112	Sequence 13112, A
30	222	13.7	175	4	US-08-311-731A-325	Sequence 325, App
31	216	13.3	524	4	US-09-252-991A-27783	Sequence 27783, A
32	213	13.1	677	4	US-09-252-991A-22442	Sequence 22442, A
33	205	12.6	370	4	US-09-134-001C-4168	Sequence 4168, Ap
34	202	12.4	367	4	US-09-489-039A-8050	Sequence 8050, Ap
35	197	12.1	521	4	US-09-543-681A-4364	Sequence 4364, Ap
36	190	11.7	525	4	US-09-424-978B-34	Sequence 34, Appl
37	170	10.5	507	4	US-09-328-352-5207	Sequence 5207, Ap
38	164	10.1	411	4	US-09-134-001C-3876	Sequence 3876, Ap
39	154.5	9.5	424	4	US-09-543-681A-4785	Sequence 4785, Ap
40	148.5	9.1	408	4	US-09-540-236-2740	Sequence 2740, Ap
41	147.5	9.1	436	3	US-08-669-378-12	Sequence 12, Appl
42	147.5	9.1	436	3	US-09-489-039A-7926	Sequence 7926, Ap
43	145.5	9.0	400	4	US-07-956-697B-5	Sequence 5, Appli
44	144.5	8.9	397	1	US-08-263-098-5	Sequence 5, Appli
45	144.5	8.9	397	1	US-08-263-098-5	Sequence 5, Appli

## ALIGNMENTS

RESULT 1  
US-09-905-290A-4  
Sequence 4, Application US/09905290A  
Patent No. 6605459  
GENERAL INFORMATION:  
APPLICANT: Rice, John  
APPLICANT: Lanning, Beth  
APPLICANT: Crawford, John  
APPLICANT: Nye, Gordon  
TITLE OF INVENTION: METHODS FOR MEASURING CYSTEINE AND DETERMINING CYSTEINE SYNTHASE  
FILE REFERENCE: Docket No. 660545920900US  
CURRENT APPLICATION NUMBER: US/09/905, 290A  
CURRENT FILING DATE: 2001-07-13  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 4  
LENGTH: 324  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-09-905-290A-4

Query Match	72.6%	Score 1179;	DB 4;	Length 324;
Best Local Similarity	68.2%	Pred. No. 7.9e-112;		
Matches 221;	Conservative 54;	Mismatches 49;	Indels 0;	Gaps 0;
QY	1	MAVERSGIAKDVTELGKTPVLYNLKADGCVARVAALKELEMPCCSVKDRIGYSMTADA	60	
DB	1	MEEDRCSIKDDATOLIGTMYLNNIVDGCVARIAKLEMEPCSSVKERLAYGMTKDA	60	
QY	61	EEKGLITPGKSVLIEPTSGNTGIGLAFAAARGYKLIITMPASMSLEIRIILAFGAELV	120	
DB	61	EDKGLITPGKSTLIBATSGNGIGLAFIGAAKGYKVALTMPSSMSLEIRIILAFGAELV	120	
QY	121	LTPDAKMAKAVOKAEELIAKTPENAYILQOFENPAMPKVHYETTGPEIWKSGDKIDAFV	180	
DB	121	LTPDSKRGVGGIIDKAEIICSKNPDSIMLEQFNPNPOTHYRTTGPEIWRDSAGEVDILV	180	
QY	181	SGIGTGTITGAKYLLKQNPNTKLGVEPVESPVLGGKPGPHKIIGIGAGFIPGYLEV	240	
DB	181	AGVGTGTLTSGSGRFLKKNKDPKVVYGVETPESAVISGKPGTHLIQIGAGLIPNDLF	240	
QY	241	NLDEVYQISSDBAIEITAKLLAKGLFVGISSGAAAAAFAQIAKRPENAGKLIVAVFES	300	
DB	241	NVDEVYQVTSVBAIEITAKLLAKGLLVGSISSGAAAAAATKVAKRPEMAGKLIVAVFES	300	
QY	301	FGERYLSVLFESVVRRAESMTFE 324		
DB	301	GGERYLSVLFESVVRRAEENLPIC 324		



Mon May 24 08:18:53 2004

us-09-931-457a-31.rai

Page 3

```

RESULT 5
US-09-489-039A-9166
; Sequence 9166, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489, 039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9166
; LENGTH: 326
; TYPE: PR1
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9166

```

Query Match	52.6%;	Score 853.5;	DB 4;	Length 326;
Best Local Similarity	54.2%;	Pred. No. 1.2e-78;		
Matches 175;	Conservative	56;	Mismatches 79;	Indels 13;
				Gaps 5;

```
QY      6 SGIAKDVTELI GKTPLVYL NKLADGCVARVA AKTLEMEPCSSVKDRIGYSMTLADAEEKGL 65
        | : : : | | | | | : : : | : : : | : : : | : : : | : : : | : : : |
Db      5 SKIFEDNSLTIGH TPLVR LNRIGNG--RI LAKVESRNPSFSVKCRIGANMIMDAEKRGV 61
QY      66 ITPGKSVLIEPTSGNTGIGLA FMAARGIKLIITMPASMSLERRIILAFGAELVLTDBA 125
        : | : | : | | | | | : : : | : : : | : : : | : : : | : : : |
Db      62 LKPGVE-LVEPTSGNTGIALAYAGARGYKTLTMPETMSTIERRKLKTIGANLVLTGBA 120
QY      126 KMKKGAVOKABEILI AKTENAV-ILQEPENPANPKVHYBTGTPEIWKGS DGIKIDAFVSGIG 184
        | | | | : | | | | : : : | | | | : : : | : : : | : : : | : : : |
Db      121 KGMKGAIQKABEIVASNP EQFLTLQESPNDANPEIHEKTGTGPEIWE DTGDGVVFISGVG 180
QY      185 TGCTITGAGKYLKEQN ENIKLI--GVBPYESPVLSGG-----KPGPHKI QGIGAGFI PG 236
        | | | : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db      181 TCGTLTGVSRYIKNTKGKCOLITAVAVEPIDSFVLAQALAGEBLKP GPHKIQGIGAGFI PG 240
QY      237 VLEVNLIDEVVQISSDEA IETAKLLALEGLFVG ISSGAAAAAFOIAKR PENNAKLI VA 296
        | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db      241 NIDLKLVDKVI GITNEEAISTARRIMEERG ILA GISSGAAVAALKQJED EAFNNKNI IV 300
QY      297 VFPSFGERYLSSVLFESVRRBAE 319
        : | | | | | : | : : | : : : | : : : | : : : | : : : |
Db      301 ILPSSGERYLSTALPADL FTEKE 323
```

RESULT 6  
US-09-107-532A-3925  
; Sequence 3925, Application US/09107532A

PILING DATE: 30-Jun-1998  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/085,598  
 FILING DATE: 14 May 1998  
 APPLICATION NUMBER: 60/051571  
 FILING DATE: July 2, 1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Ariniello, Pamela Deneke  
 REGISTRATION NUMBER: 40,489  
 REFERENCE/DOCKET NUMBER: GTC-012  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (781)893-5007  
 TELEFAX: (781)893-8277  
 INFORMATION FOR SEQ ID NO: 3925:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 308 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: YES  
 ORIGINAL SOURCE:  
 ORGANISM: Enterococcus faecium  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (B) LOCATION 1...308  
 SEQUENCE DESCRIPTION: SEQ ID NO: 3925

Query Match	50.1%;	Score 813;	DB 4;	Length 308;
Best Local Similarity	54.8%;	Pred. No. 1.5e-74;		
Matches 171;	Conservative 50;	Mismatches 85;	Indels 6;	Gaps 5;

```
QY      5 RSGIADQVTELTIGTPIPLVLNKLADGCVARVAAKLELMBECSSVKDRIGYSMIDAEK 64
        | | | | | | | | : | | | | | | | | | | | | | | | | | | | | | | |
DB      1 RDAFNSVTTELIGKTPIVKNLKI VPEBSADVFWKLEFFENOGSVKDRIALSMIEKAENDG 60
QY      65 LITPKSVLIEPTSGNTGIGAFMAAARGYKLITMPMSLSERRIILLAFGAELVLTDP 124
        | : | : | : | | | | | | : | : | : | : | : | : | : | : | : | : | : |
DB      61 LTKPGDTI-IBPTSGNTGISLWGVAKGYKVLIWMPETMSIERLLMKGYGAELLITFG 119
QY      125 AKMGCAVOKABEILAKTPNAAYILOEFENPANKVHYETGTGEIWK--GSDGKIDAFVSG 182
        | | : : : : | | | | : : | | | | | | : | : | : | : | : | : | : | : |
DB      120 ADGISGSIREAER-LAKENGYPLOLFENEANPLVHEKUTGEPEIHQAFGVNG-LDAFAVG 177
QY      183 IGTGGTTGAGKYLAKEONPNIKLIGVEPVESPVLSGKRGPBKIQIGIAGFTPGVLEVNL 242
        | | | | | | | | : | | | : | : | | | | | | | | : | : | : | : | : :
DB      178 IGTGGTTGAGRELKRVYPKIELIGVEPABSAILGEGKAGPHKIQIGITGFVPKTLDTSY 237
QY      243 LDEVVOISSDEAIETAKLLAKEGLFVGISSGAAAAAAFIQIAKPENAGKLIIVAVPESFG 302
        | : | : | | : | : | : | : | : | | | | | | | | : : | | : | : | |
DB      238 YDKVVISIGSDBAMETAREVERKERILVGISSGALIAALLKVAKE-LGKGKVALAVPDNG 296
QY      303 ERYLSSVLPESV 314
        | | | | : | : | :
DB      297 ERYSTALYOBI 308
```

RESULT 7  
US-09-134-001C-3330  
: Sequence 3330, Application US/09134001C

```

? CONDUCTOR: JAMES H. HARRIS, JR.
? ADDRESS: GENOME THERAPEUTICS CORPORATION
? STREET: 100 Beaver Street
? CITY: Waltham
? STATE: Massachusetts
? COUNTRY: USA
? ZIP: 02354
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: CD-ROM ISO9660
? COMPUTER: PC
? OPERATING SYSTEM: <Unknown>
? SOFTWARE: ASCII
?
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: HS/09/107 512A

```



SEQ ID NO 3330  
LENGTH: 319  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3330

Query Match 49.4%; Score 801; DB 4; Length 319;  
Best Local Similarity 52.6%; Pred. No. 2.7e-73;  
Matches 159; Conservative 60; Mismatches 77; Indels 6; Gaps 5;

QY 12 VTEHIGTPLYVNLKADGCVARVAAKLEMEPCSSVKDRIGYSMTADAEKGLITPGKS 71  
DB 18 VTQIIGNTPVVKLRNVVDADADIVKLEYQNPQGSVKDRIALAMIEKAEREGRIKPGDT 77  
QY 72 VLEPTSGNTGIGLAFAAARGYKLIITMPASMSLERIILAFGAELVLTDPKAMKA 131  
DB 78 I-VEPTSGNTGIGLAFAAARGYKAVFTMPETMSQERNNLLKAYGAELVLTDPSEAMKA 136  
QY 132 VQKAEILAKTPNAVYI-LOQFENPANPKVHYETTPGPEIWKSGDK-IDAFVSGIGTGTI 189  
DB 137 IKKAKEL--KEEHGFEPQGFENPANPEIHETGPELVEQFEGRIQIDAFIAGVGTGTL 194  
QY 190 TGAGKYLKEQNPNIKIGVPEVESPVLSGKPGPHKIOGIGAGFIPGVLEVNLLDEVYQI 249  
DB 195 SGVGKVLKKEYPNVEIVAIERBASPVLSGGEPPHKLQIGAGFVPDILTNTVEYDSIIKV 254  
QY 250 SSDEAIEETAKLALKEGLFVGISSGAAAAAFOIARPENAKLIVAVFPSPGRRYLSV 309  
DB 255 GNDTAMDARRVAREBEGILAGISSGAAIYAAIOKAKE-LGKGTIVTVLPISNGERYLSTP 313  
QY 310 LF 311  
DB 314 LY 315

RESULT 8

US-09-724-623-72  
Sequence 72, Application US/09724623  
Patent No. 6476209  
GENERAL INFORMATION:  
APPLICANT: Glenn, Matthew  
APPLICANT: Lubbers, Mark W  
APPLICANT: Dekker, James  
TITLE OF INVENTION: Polynucleotides, materials incorporating  
TITLE OF INVENTION: them, and methods for using them.  
FILE REFERENCE: 1048U1  
CURRENT APPLICATION NUMBER: US/09/724,623  
CURRENT FILING DATE: 2000-11-28  
NUMBER OF SEQ ID NOS: 124  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 72  
LENGTH: 309  
TYPE: PRT  
ORGANISM: Lactobacillus rhamnosus  
US-09-724-623-72

Query Match 46.9%; Score 761; DB 4; Length 309;  
Best Local Similarity 51.5%; Pred. No. 3.1e-69;  
Matches 157; Conservative 56; Mismatches 88; Indels 4; Gaps 4;

QY 9 AKDVTIELGKTPVLYNLKADGCVARVAAKLEMEPCSSVKDRIGYSMTADAEKGLITP 68  
DB 5 ADNTGLIGNTPLKLRNVVDEGADVYVLEFENPGGSVKDRIALAMIEDAEYKGLKP 64  
QY 69 GKSVLIEPTSGNTGIGLAFAAARGYKLIITMPASMSLERIILAFGAELVLTDPKAM 128  
DB 65 G-GTIVEPTSGNTGIGLAFAAARGYKLIITMPETMSQERNNLLKAYGAELVLTDPGADGM 123  
QY 129 KGAQKAEILAKTPNAVYI-LOQFENPANPKVHYETTPGPEIWKSGDKI-DAFVSGIGTGT 187  
DB 124 PGAIKKA-EALSKENGFLPMQFONPANPDVHERTTGQELIRSFDGTPDAFVAGVGTGG 182  
QY 188 TITGAGKYLKEQNPNIKIGVPEVESPVLSGKPGPHKIOGIGAGFIPGVLEVNLLDEVY 247

DB 183 TITGGERALRKINPDVQIYALAEASPMLEKGGKHKIKQISAGFIPDVLDTNLYQDI 242  
QY 248 QISSDEAIEETAKLALKEGLFVGISSGAAAAAFOIARPENAKLIVAVFPSPGRRYLS 307  
DB 243 EVTSDQALDMARHVSHEGFLPGISAGANIFGAIEILAKK-LGKGSVATVAPDNGERYLS 301  
QY 308 SVLFE 312  
DB 302 TDLFK 306

RESULT 9

US-09-543-681A-7720  
Sequence 7720, Application US/09543681A  
Patent No. 6605709  
GENERAL INFORMATION:  
APPLICANT: GARY BRETON  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRAB  
FILE REFERENCE: 2709.1002-001  
CURRENT APPLICATION NUMBER: US/09/543,681A  
CURRENT FILING DATE: 2000-04-05  
PRIOR APPLICATION NUMBER: US 60/128,706  
PRIOR FILING DATE: 1999-04-09  
NUMBER OF SEQ ID NOS: 8344  
SEQ ID NO 7720  
LENGTH: 303  
TYPE: PRT  
ORGANISM: Proteus mirabilis  
US-09-543-681A-7720

Query Match 40.1%; Score 650.5; DB 4; Length 303;  
Best Local Similarity 45.4%; Pred. No. 5.5e-58;  
Matches 138; Conservative 52; Mismatches 103; Indels 11; Gaps 4;

QY 8 IAKDVTIELGKTPVLYNLKADGCVARVAAKLEMEPCSSVKDRIGYSMTADAEKGLIT 67  
DB 10 IVAGLEQFTGNTPLVVKLQRLSGLDAETITVLEGNPNPAGSVKDRALSMIBOALRGEIK 69  
QY 68 PKSVLIEPTSGNTGIGLAFAAARGYKLIITMPASMSLERIILAFGAELVLTDPKAM 127  
DB 70 PG-DTILRATSGNTGIGLAFAAARGYKLIITMPENMSKERQASMAAGAEILVSRBIG 128  
QY 128 MKGAVQKAEILAKTPNAVYI-LOQFENPANPKVHYETTPGPEIWKSGDKIDAFVSGIGTGT 187  
DB 129 MEGARDIAQF-MERKGEKVLDOFNPNPRAHPTSTGPEIMQOTGRITHFVSSMGTTG 187  
QY 188 TITGAGKYLKEQNPNIKIGVPEVESPVLSGKPGPHKIOGIGAGFIPGVLEVNLLDEVY 247  
DB 188 TITGVSYLKTQSDTVKIVGLQPEKSGIPIG-----IRWSPAYLPGIFRKELVDSVI 240  
QY 248 QISSDEAIEETAKLALKEGLFVGISSGAAAAAFOIARPENAKLIVAVFPSPGRRYLS 307  
DB 241 DMSQTBAGQITWELLASKGIFCGVSSGGAAGAIRVAK--ENPGAVIYVAIICDRGRYLS 298  
QY 308 SVLFE 311  
DB 299 TGVF 302

RESULT 10

US-09-489-039A-9175  
Sequence 9175, Application US/09489039A  
Patent No. 6610836  
GENERAL INFORMATION:  
APPLICANT: Gary Breton et. al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
FILE REFERENCE: 2709.2004001  
CURRENT APPLICATION NUMBER: US/09/489,039A  
CURRENT FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: US 60/117,747

PRIOR FILING DATE: 1999-01-29  
 ; NUMBER OF SEQ ID NOS: 14342  
 ; SEQ ID NO 9175  
 ; LENGTH: 306  
 ; TYPE: PRT  
 ; ORGANISM: Klebsiella pneumoniae  
 US-09-489-039A-9175

Query Match 38.7%; Score 628.5; DB 4; Length 306;  
 Best Local Similarity 44.4%; Pred. No. 9.8e-56;  
 Matches 135; Conservative 56; Mismatches 102; Indels 11; Gaps 4;

QY 8 IAKDVTEIGKTPLVYLNKLDGCVARVAAKLEMEPCSSVKDRIGYSMIADAEKGLIT 67  
 DB 3 IVNTLEQITGNTPLVKLQRLGPDNGSEVWVKLEGNPNAGSVKDRALSMITVEAEKRGETQ 62  
 QY 68 PEKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILLAFGAELVLTDPAKG 127  
 DB 63 PG-DVLEATSGNTGIALAMIAALKYRMKLLMPDNMSQERRAAMRAYGAELILVSKEQG 121  
 QY 128 MKGAVOKAEELIAKTPNAVYILQGFENPANKVHYETTGPEIWKSGDKIDAFVSGIGTGG 187  
 DB 122 MEGARDLAE-MAORGEGKLEDPNNPDNPYHYTTGTPEIWOQTAGRITHFVSSMGTTG 180  
 QY 188 TITGAGKYLKEQNPNIKLGIVEPVESPVLSGSGPKHIGIGAGFIPGVLEVNLDDEV 247  
 DB 181 TITGVSRLREKSPVTIVGLQPEGSSIPGIRRP-----AEYMPGIFNASTVDTVL 233  
 QY 248 QISDRAIETAKLALKEGLFVGISGAAAFQIAKRPENAGKLIIVAFPSFGERYLS 307  
 DB 234 DIHQDAENTMRQLAVREGIFCGVSSGAAVAGALRIAR--ENPGAVVAIVCDRGDRYLS 291  
 QY 308 SVLF 311  
 DB 292 TGVF 295

RESULT 11  
 US-09-252-991A-29566  
 ; Sequence 29566, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252.991A  
 ; PRIOR APPLICATION NUMBER: 1999-02-18  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094.190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 29566  
 ; LENGTH: 382  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-29566

Query Match 36.9%; Score 598.5; DB 4; Length 382;  
 Best Local Similarity 41.9%; Pred. No. 1.6e-52;  
 Matches 126; Conservative 60; Mismatches 104; Indels 11; Gaps 4;

QY 12 VTELIKTPVLYLNKLDGCVARVAAKLEMEPCSSVKDRIGYSMIADAEKGLITPKGS 71  
 DB 91 IADCVGNTPLVKLQRLGPDNGSEVWVKLEGNPNAGSVKDRALSMITVEAEKRGETQ 149  
 QY 72 VLIBPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILLAFGAELVLTDPAKMGKA 131  
 DB 150 TLIBATSGNTGIALAMIAALKYRMKLLMPDNMSQERRAAMRAYGAELILVSKEGMEGA 209  
 QY 132 VQAEELIAKTPNAVYILQGFENPANKVHYETTGPEIWKSGDKIDAFVSGIGTGTITG 191

DB 210 RDLADK-LQREGNGKVLDPFANGDNPEAHYHSTGPEIWOQTGSIHFVSSMGTTGTMG 268  
 QY 192 AGKYLKEQNPNIKLGIVEPVESPVLSGSGPKHIGIGAGFIPGVLEVNLDDEVQISS 251  
 DB 269 VSRYLKEQNPNAVIGLQPMESGALPGIRMPQE-----YLPKIYDASRVDRVDMHQ 321  
 QY 252 DEAIETAKLALKEGLFVGISGAAAFQIAKRPENAGKLIIVAFPSFGERYLSVLF 311  
 DB 322 DEADIMRRLAREGIFCGVSSGAAVAAWLRLSRELENA--VLVAICDRGDRYLSGVY 379  
 QY 312 E 312  
 DB 380 D 380

RESULT 12  
 US-09-328-352-8179  
 ; Sequence 8179, Application US/09328352  
 ; Patent No. 6562958  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gary L. Breton et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
 ; FILE REFERENCE: GTC99-03PA  
 ; CURRENT APPLICATION NUMBER: US/09/328.352  
 ; CURRENT FILING DATE: 1999-06-04  
 ; NUMBER OF SEQ ID NOS: 8252  
 ; SEQ ID NO 8179  
 ; LENGTH: 307  
 ; TYPE: PRT  
 ; ORGANISM: Acinetobacter baumannii  
 US-09-328-352-8179

Query Match 36.6%; Score 594; DB 4; Length 307;  
 Best Local Similarity 45.9%; Pred. No. 3.3e-52;  
 Matches 136; Conservative 49; Mismatches 97; Indels 14; Gaps 5;

QY 16 IGTPLVLYLNKLDGCVARVAAKLEMEPCSSVKDRIGYSMIADAEKGLITPKGSVLE 75  
 DB 22 VGTPLVLYLNKLDGCVARVAAKLEMEPCSSVKDRIGYSMIADAEKGLITPKGSVLE 80  
 QY 76 PTSNTGIGLAFMAAARGYKLIITMPASMSLERRIILLAFGAELVLTDPAKMGAGVOKA 135  
 DB 81 ATSGNTGIALAMIAALKYRMKLLMPDNMSQERRAAMRAYGAELI--EAPNMEARDMA 137  
 QY 136 BRIIAKTPNAVYILQGFENPANKVHYETTGPEIWKSGDKIDAFVSGIGTGTITGAGKY 195  
 DB 138 LQMOAE-GLGLVINOFPNDVNEAHYLTGPEIMQGTGKITHFVSSMGTTGTMGSKY 196  
 QY 196 LKQNPNIKLGIVEPVESPVLSGSGPKHIGIGAGFIPGVLEVNLDDEVQISSDEAI 255  
 DB 197 LKQNPDIQIIGLQPSGNSINAGIRMPQE-----YLPTEFEPKRVQIMDIPQTEAE 249  
 QY 256 ETAKLALKEGLFVGISGAAAFQIAKRPENAGKLIIVAFPSFGERYLSVLF 311  
 DB 250 KTAIRARREGISAGTSSGAAVAAWLRLSRELENA--ENPDVAVICICDRGDRYLSGLF 303

RESULT 13  
 US-09-540-236-3342  
 ; Sequence 3342, Application US/09540236  
 ; Patent No. 6673910  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gary L. Breton et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATF  
 ; FILE REFERENCE: 2709.2005-001  
 ; CURRENT APPLICATION NUMBER: US/09/540.236  
 ; CURRENT FILING DATE: 2000-04-04  
 ; NUMBER OF SEQ ID NOS: 3840  
 ; SEQ ID NO 3342  
 ; LENGTH: 310  
 ; TYPE: PRT

ORGANISM: M.catarhalis  
US-09-540-236-3342

Query Match 36.5%; Score 593; DB 4; Length 310;  
Best Local Similarity 42.5%; Pred. No. 4.2e-52;  
Matches 135; Conservative 58; Mismatches 105; Indels 20; Gaps 6;

QY 1 MAVERSGIARD-----VTELIKTPLYENKLDGCVARVAAKLEMEPCSSVKDRIG 53  
Db 2 MIYETTTMTKDFHQTTLLSDCVGNTPLVKNLTLAKDSQATLLAKLEGNPFAGSVKDRPA 61  
QY 54 YSMIADAEKGLTFPGKSVLIEPTSGNTGIGLAFMAARGYKLIITMPASMSLERIILL 113  
Db 62 FNMIVQAEQGRQIKPG-DTLIBATSGNTGIALMAVAMRGYQMTLMPANSTQERKDMT 120  
QY 114 AFGABVLVTDPAKMGKAGVQKABEILAKTPNAVILQOFENPANPKVHYETGTGPEIWKGS 173  
Db 121 AYGATLI--EVADGMAARMDALQMA-DGGGIVLDQFNNTDNKHAHYLTGPBLMQOTE 177  
QY 174 GKIDAFVSGIGTGTITGAGKYLKEQNPNIKLGVEPVESPVLSGKPGPKRIGIGAGF 233  
Db 178 GKITHFVSSMGTTGTIMGVSCYLKEQNPALQIGLPDEQSSIAIGIRRP-----ABY 230  
QY 234 IPGVLEVNLDDEVQISSDEALETAKLLAKKEGLFVGISSGAAAAAFQIAKRPENAKL 293  
Db 231 LFGIPDASLVDTIMDVQNTAERYMKLAREEGIFCGVSSGGAAMAHAQIAKNTPD--V 288  
QY 294 IVAVFPSPFGERYLSSVLF 311  
Db 289 IAFIVCDRGDRYLSTGLF 306

RESULT 14

US-09-789-300A-4  
Sequence 4, Application US/09789300A  
Patent No. 6458576  
GENERAL INFORMATION:  
APPLICANT: Meyers, Rachel  
APPLICANT: Rudolph-Owen, Laura A.  
TITLE OF INVENTION: 22406, A No. 6458576el Human Pyridoxal-Phosphate  
TITLE OF INVENTION: Dependent Enzyme Family Member and Uses Therefore  
FILE REFERENCE: 35800/208926  
CURRENT APPLICATION NUMBER: US/09/789,300A  
CURRENT FILING DATE: 2001-02-20  
PRIOR APPLICATION NUMBER: US 60/163,208  
PRIOR FILING DATE: 2000-02-17  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 378  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURES:  
OTHER INFORMATION: Pyridoxal-Phosphate Dependent Enzyme Family Domain  
OTHER INFORMATION: Sequence  
US-09-789-300A-4

Query Match 36.0%; Score 584; DB 4; Length 378;  
Best Local Similarity 43.0%; Pred. No. 4.7e-51;  
Matches 161; Conservative 45; Mismatches 80; Indels 88; Gaps 17;

QY 12 VTELIKTPLYENKLA-----DGCVARVAAKL-----ELMEPCSSVKDR-IGYSMIADAE 61  
Db 1 VTELIKTPLYENKLA-----DGCVARVAAKL-----ELMEPCSSVKDR-IGYSMIADAE 60  
QY 62 ---EKGLITRG-----KSVLEPTSGNTGIGLAFMAARGYKLIITMPAS-MSLERIIT 111  
Db 61 KLGKKGIVPGTVQVESKTIIEPTSGNTGIALMAALGLKCTIVMPATDTSEKTAQ 120  
QY 112 LEAFGAEVLVTDPAK-----XKGAVQKABEILAKTP-NAVILQO---FENPANPKV-HY 161  
Db 121 LEAFGAEVLVTDPAK-----XKGAVQKABEILAKTP-NAVILQO---FENPANPKV-HY 160

QY 162 ETTGPEIWKSGDGK-----IDAFVSGIGTGTITGAGKYLKEQNP-----IKL 205  
Db 181 KTIGPRIMEQLGKEISIGRLPDVAVAVGGGTTTGARLYKELNPDGKIDVLEPVAV 240  
QY 206 IGVPEVSPVLG-----GKPGP-----HKIQIGAG----- 232  
Db 241 IGVPEVSPVLG-----GKPGP-----HKIQIGAG----- 230  
QY 233 -FIPGVLEVNLDDEVQISSDEALETAKLLAKKEGLFVGISSGAAAAAFQIAK-- 284  
Db 301 EFVPRILDELDRQIGDEVVTVTDEALERARLARBEGLIVGSSGAAVAAALKLAKE 360  
QY 285 -KRPENAGKLIIVAV 297  
Db 361 GKPLNKGKTIIVVI 374

RESULT 15

US-09-107-532A-6629  
Sequence 6629, Application US/09107532A  
Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: LYNN A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESS: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 6629:

SEQUENCE CHARACTERISTICS:  
LENGTH: 308 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (B) LOCATION 1...308

SEQUENCE DESCRIPTION: SEQ ID NO: 6629:

US-09-107-532A-6629

Query Match 34.0%; Score 552.5; DB 4; Length 308;  
Best Local Similarity 38.8%; Pred. No. 5.6e-48;  
Matches 120; Conservative 65; Mismatches 113; Indels 11; Gaps 6;

QY 8 TARDYTELIGTPLYTN-----KLADGCVARVAAKLEMEPCSSVKDRIGYSMIADAEK 63



Db 5 IISTITTEAIGSTPLLYRFDNQHYSMRPG--SAIYAKLEIYLNPGSSIKDRLGSYLIKEGFAQ 62  
QY 64 GLITPGKSVLIEPTSGNTGIGLAFMAARGYKLIITMPASMSLEIRIILLAFGAEIVLTD 123  
Db 63 GKINE-KTIIIEPTAGNTGIGLALALTYKLRVFWPEKPSLEKQOLMKGAKIHSP 121  
QY 124 PAKMGAVQKABEILAKTPNAVYILQFENPANPKVHYETTPPEIWKSGSDKIDAFVSGI 183  
Db 122 SEQIGIGAISSKSKLAERISNSYLPLOFENKNDPAAYYHTLGPETIPEIKENIHSEVAGI 181  
QY 184 GTGGTITGAGKYLKQNPNIKEIGVEPVESPVLSGKPGPHKIQIGAGFTPGVLEVNLL 243  
Db 182 GSGGTFAGTSTFLKPKYDRIIGVQ-EGSVLNGGDPAPHEIEGIGVEFIPLPLSLST 240  
QY 244 DEWQISSDEAIEETAKLLALKEGLFVGISGAAAAAF-QIAKRENAKGLIVAVFPSPG 302  
Db 241 NQIETISDVEGFNYTRQLABEQGLVGSIGAAFAAALREIRRLP--PGHRVVTIFPDAA 298  
QY 303 ERYLSSVLE 311  
Db 299 DRYLSKNY 307

Search completed: May 18, 2004, 13:09:40  
Job time : 24 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2004, 13:08:40 ; Search time 48 Seconds  
(without alignments)  
1884.062 Million cell updates/sec

Title: US-09-931-457A-31  
Perfect score: 1623  
Sequence: 1 MAVERSGIAKDVTTELIGTTP.....LSSVLFESVREAESMTFRP 325

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1145568 segs, 278261457 residues  
Total number of hits satisfying chosen parameters: 1145568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1623	100.0	325	9 US-09-931-457A-31	Sequence 31, Appl
2	1622	99.9	325	12 US-10-424-599-273302	Sequence 273302,
3	1488	91.7	325	12 US-10-424-599-198006	Sequence 198006,
4	1453	89.5	325	9 US-09-931-457A-32	Sequence 32, Appl
5	1389	85.6	328	12 US-10-425-114-59830	Sequence 59830, A
6	1381	85.1	328	12 US-10-425-114-47602	Sequence 47602, A
7	1381	85.1	328	12 US-10-425-114-49092	Sequence 49092, A
8	1380	85.0	328	12 US-10-425-114-59613	Sequence 59613, A
9	1375	84.7	328	12 US-10-425-114-60606	Sequence 60606, A
10	1364	84.0	325	12 US-10-464-811-21	Sequence 21, Appl
11	1317	81.1	415	9 US-09-931-457A-64	Sequence 64, Appl
12	1314	81.0	386	9 US-09-931-457A-66	Sequence 66, Appl
13	1309	80.7	316	12 US-10-425-114-60706	Sequence 60706, A
14	1279	78.8	397	12 US-10-425-114-57303	Sequence 57303, A
15	1279	78.8	398	12 US-10-425-114-71564	Sequence 71564, A

16	1275	78.6	398	9 US-09-931-457A-62	Sequence 62, Appl
17	1275	78.6	402	12 US-10-425-114-69297	Sequence 69297, A
18	1271	78.3	383	9 US-09-931-457A-65	Sequence 65, Appl
19	1234	76.0	324	12 US-10-424-599-204076	Sequence 204076,
20	1231	75.8	290	12 US-10-424-599-198002	Sequence 198002,
21	1179	72.6	324	10 US-09-905-290A-4	Sequence 4, Appl
22	1079	66.5	372	12 US-10-424-599-196219	Sequence 196219,
23	985	60.7	212	12 US-10-425-114-45564	Sequence 45564, A
24	968.5	59.7	315	12 US-10-282-122A-48420	Sequence 48420, A
25	949.5	58.5	310	12 US-10-282-122A-62704	Sequence 62704, A
26	949.5	58.4	310	12 US-10-282-122A-64692	Sequence 64692, A
27	947.5	58.4	328	12 US-10-424-599-160879	Sequence 160879,
28	941.5	58.0	310	12 US-10-282-122A-62078	Sequence 62078, A
29	940.5	57.9	310	12 US-10-282-122A-63762	Sequence 63762, A
30	917	56.5	309	12 US-10-282-122A-63092	Sequence 63092, A
31	916.5	56.5	309	12 US-10-282-122A-52203	Sequence 52203, A
32	912	56.2	310	12 US-10-282-122A-65838	Sequence 65838, A
33	896.5	55.2	309	12 US-10-282-122A-52032	Sequence 52032, A
34	882	54.3	322	12 US-10-282-122A-77194	Sequence 77194, A
35	879	54.2	317	12 US-10-282-122A-68735	Sequence 68735, A
36	870	53.6	332	12 US-10-282-122A-45004	Sequence 45004, A
37	865.5	53.3	323	12 US-10-282-122A-74982	Sequence 74982, A
38	865.5	53.3	323	12 US-10-282-122A-75586	Sequence 75586, A
39	865.5	53.3	323	12 US-10-464-811-22	Sequence 22, Appl
40	863.5	53.2	323	12 US-10-282-122A-55655	Sequence 55655, A
41	862	53.1	322	12 US-10-282-122A-77839	Sequence 77839, A
42	858.5	52.9	323	9 US-09-815-242-13862	Sequence 13862, A
43	857.5	52.8	323	9 US-09-815-242-10239	Sequence 10239, A
44	857.5	52.8	323	12 US-10-282-122A-56629	Sequence 56629, A
45	856	52.7	324	9 US-09-815-242-11879	Sequence 11879, A

ALIGNMENTS

RESULT 1  
US-09-931-457A-31  
; Sequence 31, Application US/09931457A  
; Patent No. US20020157132A1  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; TITLE OF INVENTION: Plant Amino Acid Biosynthetic Enzymes  
; FILE REFERENCE: B8116 US CIP  
; CURRENT APPLICATION NUMBER: US/09/931,457A  
; PRIOR FILING DATE: 2002-02-22  
; PRIOR APPLICATION NUMBER: 09/424, 976  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: 60/065, 385  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/049, 406  
; PRIOR FILING DATE: 1997-06-12  
; NUMBER OF SEQ ID NOS: 72  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 31  
; LENGTH: 325  
; TYPE: PRT  
; ORGANISM: Glycine max  
US-09-931-457A-31

Query Match 100.0%; Score 1623; DB 9; Length 325;  
Best Local Similarity 100.0%; Pred. No. 1.1e-154;  
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAVERSGIAKDVTTELIGTTPVLVYLKADGCVARYAAKLEIMBPCSSVKDRIGYSMTADA	60
Db	1	MAVERSGIAKDVTTELIGTTPVLVYLKADGCVARYAAKLEIMBPCSSVKDRIGYSMTADA	60
Qy	61	EEKGLITPGKSVLIBPTSGNTGIGLAFFMAARGYKLIITMPASMSLERRIILAFGAEIV	120
Db	61	EEKGLITPGKSVLIBPTSGNTGIGLAFFMAARGYKLIITMPASMSLERRIILAFGAEIV	120
Qy	121	LTPDAKPKGAVQKAEELIAKTPNAYIIQQFENPANPKVHYETTGPEIMKGSDDKIDAFV	180

```

Db      121 LTPAKGMKGAQKAEIILAKTPNAYILQGFENPANPKVHYETTGPETWKGSDGKIDAFV 180
      181 SGIGTGTITGAGKYLKEQNPNIKIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEV 240
      181 SGIGTGTITGAGKYLKEQNPNIKIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEV 240
QY      241 NLDEVVOISSDEAIFETAKLALKEGLFVGISGAAAAAFOIAKRPENAGKLIIVAVFPS 300
      241 NLDEVVOISSDEAIFETAKLALKEGLFVGISGAAAAAFOIAKRPENAGKLIIVAVFPS 300
Db      301 FGERYLSSVLPESVRRBAESMTPEP 325
      301 FGERYLSSVLPESVRRBAESMTPEP 325

```

RESULT 2

```

US-10-424-599-273302
; Sequence 273302, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 273302
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_88813C.1.pep
US-10-424-599-273302

```

Query Match 99.9%; Score 1622; DB 12; Length 325;  
 Best Local Similarity 99.7%; Pred. No. 1.4e-154;  
 Matches 324; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 MAVERSGIAKDVTELIGKTPLYLNLKADGCVARVAKLELMEPCSSVKDRIGYSMTADA 60
      1 MAVERSGIAKDVTELIGKTPLYLNLKADGCVARVAKLELMEPCSSVKDRIGYSMTADA 60
Db      61 SEKGLITPKSVLIBPTSGNTGICLAFMAAARGYKLIITMPASMSLERRITLLAFGAEIV 120
      61 SEKGLITPKSVLIBPTSGNTGICLAFMAAARGYKLIITMPASMSLERRITLLAFGAEIV 120
QY      121 LTPAKGMKGAQKAEIILAKTPNAYILQGFENPANPKVHYETTGPETWKGSDGKIDAFV 180
      121 LTPAKGMKGAQKAEIILAKTPNAYILQGFENPANPKVHYETTGPETWKGSDGKIDAFV 180
Db      121 LTPAKGMKGAQKAEIILAKTPNAYILQGFENPANPKVHYETTGPETWKGSDGKIDAFV 180
      121 LTPAKGMKGAQKAEIILAKTPNAYILQGFENPANPKVHYETTGPETWKGSDGKIDAFV 180
QY      181 SGIGTGTITGAGKYLKEQNPNIKIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEV 240
      181 SGIGTGTITGAGKYLKEQNPNIKIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEV 240
Db      181 SGIGTGTITGAGKYLKEQNPNIKIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEV 240
      181 SGIGTGTITGAGKYLKEQNPNIKIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEV 240
QY      241 NLDEVVOISSDEAIFETAKLALKEGLFVGISGAAAAAFOIAKRPENAGKLIIVAVFPS 300
      241 NLDEVVOISSDEAIFETAKLALKEGLFVGISGAAAAAFOIAKRPENAGKLIIVAVFPS 300
Db      241 NLDEVVOISSDEAIFETAKLALKEGLFVGISGAAAAAFOIAKRPENAGKLIIVAVFPS 300
      241 NLDEVVOISSDEAIFETAKLALKEGLFVGISGAAAAAFOIAKRPENAGKLIIVAVFPS 300
QY      301 FGERYLSSVLPESVRRBAESMTPEP 325
      301 FGERYLSSVLPESVRRBAESMTPEP 325
Db      301 FGERYLSSVLPESVRRBAESMTPEP 325
      301 FGERYLSSVLPESVRRBAESMTPEP 325

```

RESULT 3  
 US-10-424-599-198006  
 ; Sequence 198006, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:

```

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 198006
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_20824C.1.pep
US-10-424-599-198006

```

Query Match 91.7%; Score 1488; DB 12; Length 325;  
 Best Local Similarity 89.2%; Pred. No. 4.3e-141;  
 Matches 290; Conservative 21; Mismatches 14; Indels 0; Gaps 0;

```

QY      1 MAVERSGIAKDVTELIGKTPLYLNLKADGCVARVAKLELMEPCSSVKDRIGYSMTADA 60
      1 MAVERSGIAKDVTELIGKTPLYLNLKADGCVARVAKLELMEPCSSVKDRIGYSMTADA 60
Db      61 SEKGLITPKSVLIBPTSGNTGICLAFMAAARGYKLIITMPASMSLERRITLLAFGAEIV 120
      61 SEKGLITPKSVLIBPTSGNTGICLAFMAAARGYKLIITMPASMSLERRITLLAFGAEIV 120
QY      121 LTPAKGMKGAQKAEIILAKTPNAYILQGFENPANPKVHYETTGPETWKGSDGKIDAFV 180
      121 LTPAKGMKGAQKAEIILAKTPNAYILQGFENPANPKVHYETTGPETWKGSDGKIDAFV 180
Db      121 LTPAKGMKGAQKAEIILAKTPNAYILQGFENPANPKVHYETTGPETWKGSDGKIDAFV 180
      121 LTPAKGMKGAQKAEIILAKTPNAYILQGFENPANPKVHYETTGPETWKGSDGKIDAFV 180
QY      181 SGIGTGTITGAGKYLKEQNPNIKIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEV 240
      181 SGIGTGTITGAGKYLKEQNPNIKIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEV 240
Db      181 SGIGTGTITGAGKYLKEQNPNIKIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEV 240
      181 SGIGTGTITGAGKYLKEQNPNIKIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEV 240
QY      241 NLDEVVOISSDEAIFETAKLALKEGLFVGISGAAAAAFOIAKRPENAGKLIIVAVFPS 300
      241 NLDEVVOISSDEAIFETAKLALKEGLFVGISGAAAAAFOIAKRPENAGKLIIVAVFPS 300
Db      241 NLDEVVOISSDEAIFETAKLALKEGLFVGISGAAAAAFOIAKRPENAGKLIIVAVFPS 300
      241 NLDEVVOISSDEAIFETAKLALKEGLFVGISGAAAAAFOIAKRPENAGKLIIVAVFPS 300
QY      301 FGERYLSSVLPESVRRBAESMTPEP 325
      301 FGERYLSSVLPESVRRBAESMTPEP 325
Db      301 FGERYLSSVLPESVRRBAESMTPEP 325
      301 FGERYLSSVLPESVRRBAESMTPEP 325

```

RESULT 4  
 US-09-931-457A-32  
 ; Sequence 32, Application US/09931457A  
 ; Patent No. US20020157132A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Falco, S. Carl  
 ; APPLICANT: Allen, Stephen M.  
 ; TITLE OF INVENTION: Plant Amino Acid Biosynthetic Enzymes  
 ; FILE REFERENCE: B81116 US CIP  
 ; CURRENT APPLICATION NUMBER: US/09/931,457A  
 ; CURRENT FILING DATE: 2002-02-22  
 ; PRIOR APPLICATION NUMBER: 09/424,976  
 ; PRIOR FILING DATE: 1999-12-02  
 ; PRIOR APPLICATION NUMBER: 60/065,385  
 ; PRIOR FILING DATE: 1997-11-12  
 ; PRIOR APPLICATION NUMBER: 60/049,406  
 ; PRIOR FILING DATE: 1997-06-12  
 ; NUMBER OF SEQ ID NOS: 72  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 32  
 ; LENGTH: 325  
 ; TYPE: PRT  
 ; ORGANISM: Citrullus lanatus  
 US-09-931-457A-32  
 Query Match 89.5%; Score 1453; DB 9; Length 325;



Best Local Similarity 87.1%; Pred. No. 1.5e-137;  
Matches 283; Conservative 24; Mismatches 18; Indels 0; Gaps 0;

QY 1 MAVERSGIAKDVTELGKTPVLYLNKLADGCVARVAAKLEMEPCSSVKDRIGYSMTADA 60  
DB 1 MADAKSTIAKDVTELGKTPVLYLNKLADGCVARVAAKLEMEPCSSVKDRIGYSMTADA 60  
QY 61 EEKGLITPCKSVLIEPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILAFGAELV 120  
DB 61 ENKGLITPCKSVLIEPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILAFGAELV 120  
QY 121 LTDPKAKMGAVOKAEIILAKTPNAYILLOQFENPANPKHYETTGTPEIWKSGDKIDAFV 180  
DB 121 LTDPKAKMGAVOKAEIILAKTPNAYILLOQFENPANPKHYETTGTPEIWKSGDKIDAFV 180  
QY 181 SGITGCTITGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEV 240  
DB 181 SGITGCTITGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEV 240  
QY 241 NLDEVVQISSDEAIEITAKLALKEGLFVGISSGAAAAAFQIAKRPENAGKLIYAVFPS 300  
DB 241 NLDEVVQISSDEAIEITAKLALKEGLFVGISSGAAAAAFQIAKRPENAGKLIYAVFPS 300  
QY 301 FGERYLSVLFESVRRREASMTFEP 325  
DB 301 FGERYLSVLFESVRRREASMTFEP 325

RESULT 5  
US-10-425-114-59830  
Sequence 59830, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:

APPLICANT: Liu, Jingdong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E  
APPLICANT: Tabaska, Jack E  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 59830  
LENGTH: 328  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB3900-005-C7\_FLI.pep  
US-10-425-114-59830

Query Match 85.6%; Score 1389; DB 12; Length 328;  
Best Local Similarity 84.9%; Pred. No. 4.2e-131;  
Matches 270; Conservative 22; Mismatches 26; Indels 0; Gaps 0;

QY 8 IAKDVTELGKTPVLYLNKLADGCVARVAAKLEMEPCSSVKDRIGYSMTADAEEKGLIT 67  
DB 11 IAKDVTELGKTPVLYLNKLADGCVARVAAKLEMEPCSSVKDRIGYSMTADAEEKGLIT 70  
QY 68 PGKSVLIBPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILAFGAELVLTDPKAG 127  
DB 71 PGKSVLIBPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILAFGAELVLTDPKAG 130  
QY 128 MKGAVOKAEIILAKTPNAYILLOQFENPANPKHYETTGTPEIWKSGDKIDAFVSGITGG 187  
DB 131 MKGAVOKAEIILAKTPNAYILLOQFENPANPKHYETTGTPEIWKSGDKIDAFVSGITGG 190  
QY 188 TITGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEVNLLDEV 247  
DB 191 TITGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEVNLLDEV 250

QY 248 QISSDEAIEITAKLALKEGLFVGISSGAAAAAFQIAKRPENAGKLIYAVFPSGERYLS 307  
DB 251 QISSDEAIEITAKLALKEGLFVGISSGAAAAAFQIAKRPENAGKLIYAVFPSGERYLS 310  
QY 308 SVLFESVRRREASMTFEP 325  
DB 311 SVLFESVRRREASMTFEP 328

RESULT 6  
US-10-425-114-47602  
Sequence 47602, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:

APPLICANT: Liu, Jingdong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E  
APPLICANT: Tabaska, Jack E  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 47602  
LENGTH: 328  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: 700153393\_FLI.pep  
US-10-425-114-47602

Query Match 85.1%; Score 1381; DB 12; Length 328;  
Best Local Similarity 84.6%; Pred. No. 2.7e-130;  
Matches 269; Conservative 22; Mismatches 27; Indels 0; Gaps 0;

QY 8 IAKDVTELGKTPVLYLNKLADGCVARVAAKLEMEPCSSVKDRIGYSMTADAEEKGLIT 67  
DB 11 IAKDVTELGKTPVLYLNKLADGCVARVAAKLEMEPCSSVKDRIGYSMTADAEEKGLIT 70  
QY 68 PGKSVLIBPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILAFGAELVLTDPKAG 127  
DB 71 PGKSVLIBPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILAFGAELVLTDPKAG 130  
QY 128 MKGAVOKAEIILAKTPNAYILLOQFENPANPKHYETTGTPEIWKSGDKIDAFVSGITGG 187  
DB 131 MKGAVOKAEIILAKTPNAYILLOQFENPANPKHYETTGTPEIWKSGDKIDAFVSGITGG 190  
QY 188 TITGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEVNLLDEV 247  
DB 191 TITGAGKYLKEQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEVNLLDEV 250  
QY 248 QISSDEAIEITAKLALKEGLFVGISSGAAAAAFQIAKRPENAGKLIYAVFPSGERYLS 307  
DB 251 QISSDEAIEITAKLALKEGLFVGISSGAAAAAFQIAKRPENAGKLIYAVFPSGERYLS 310  
QY 308 SVLFESVRRREASMTFEP 325  
DB 311 SVLFESVRRREASMTFEP 328

RESULT 7  
US-10-425-114-49092  
Sequence 49092, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jingdong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E  
APPLICANT: Tabaska, Jack E

APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 49092  
LENGTH: 328  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB3150-027-D8\_Fli.pep  
US-10-425-114-49092

Query Match 85.1%; Score 1381; DB 12; Length 328;  
Best Local Similarity 84.6%; Pred. No. 2.7e-130;  
Matches 269; Conservative 22; Mismatches 27; Indels 0; Gaps 0;

QY 8 IAKDVTLLIGKTPVLYLNKLDGCVARVAAKLELMEPCSSVKDRIGYSMTADAEKGLIT 67  
DB 11 IAKDVTLLIGKTPVLYLNKLDGCVARVAAKLELMEPCSSVKDRIGYSMTADAEKGLIT 70  
QY 68 PKSVLIBPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILAFGAELVLTDPAG 127  
DB 71 PGVSVLIBPTSGNTGIGLAFAAARGYKLIITMPASMSMERRIILAFGAELVLTDPAG 130  
QY 128 MKGAVOKAEIILAKTPNAYILQGFENPANKVHYETTGPEIWKSGDKIDAFVSGIGTGG 187  
DB 131 MKGAVOKAEIILAKTPNAYILQGFENPANKVHYETTGPEIWKATAGKIDAFVSGIGTGG 190  
QY 188 TITGAGKYLKEONPNKILIGVEPVESVLSGKPGPHKIQIGAGFIPGVLEVNLLDEVV 247  
DB 191 TITGGRYLRQNPVNLKYGVEPVESAVLNGKPGPHKIQIGAGFIPGVLDVLLDETL 250  
QY 248 QISSDEALETAKLALKEGLFVGISGAAAAAFQIAKRPENAGKLIIVAVPSPGGRYLS 307  
DB 251 QVSSDEALETAKLALKEGLLVGISSGAAAAAVRLAKRPENAGKLIIVAVPSPGGRYLS 310  
QY 308 SVLFESVREAESMTFEP 325  
DB 311 SVLFQSIKKEAESMTVEP 328

RESULT 8  
US-10-425-114-59613  
Sequence 59613, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jindong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E.  
APPLICANT: Tabaska, Jack E.  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 59613  
LENGTH: 328  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB3597-017-F4\_Fli.pep  
US-10-425-114-59613

Query Match 85.0%; Score 1380; DB 12; Length 328;  
Best Local Similarity 84.6%; Pred. No. 3.4e-130;  
Matches 269; Conservative 22; Mismatches 27; Indels 0; Gaps 0;

QY 8 IAKDVTLLIGKTPVLYLNKLDGCVARVAAKLELMEPCSSVKDRIGYSMTADAEKGLIT 67  
DB 11 IAKDVTLLIGKTPVLYLNKLDGCVARVAAKLELMEPCSSVKDRIGYSMTADAEKGLIT 70  
QY 68 PKSVLIBPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILAFGAELVLTDPAG 127  
DB 71 PGVSVLIBPTSGNTGIGLAFAAARGYKLIITMPASMSMERRIILAFGAELVLTDPAG 130  
QY 128 MKGAVOKAEIILAKTPNAYILQGFENPANKVHYETTGPEIWKSGDKIDAFVSGIGTGG 187  
DB 131 MKGAVOKAEIILAKTPNAYILQGFENPANKVHYETTGPEIWKATAGKIDAFVSGIGTGG 190  
QY 188 TITGAGKYLKEONPNKILIGVEPVESVLSGKPGPHKIQIGAGFIPGVLEVNLLDEVV 247  
DB 191 TITGGRYLRQNPVNLKYGVEPVESAVLNGKPGPHKIQIGAGFIPGVLDVLLDETL 250  
QY 248 QISSDEALETAKLALKEGLFVGISGAAAAAFQIAKRPENAGKLIIVAVPSPGGRYLS 307  
DB 251 QVSSDEALETAKLALKEGLLVGISSGAAAAAVRLAKRPENAGKLIIVAVPSPGGRYLS 310  
QY 308 SVLFESVREAESMTFEP 325  
DB 311 SVLFQSIKKEAESMTVEP 328

RESULT 9  
US-10-425-114-60606  
Sequence 60606, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jindong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E.  
APPLICANT: Tabaska, Jack E.  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 60606  
LENGTH: 328  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB3597-067-G10\_Fli.pep  
US-10-425-114-60606

Query Match 84.7%; Score 1375; DB 12; Length 328;  
Best Local Similarity 84.3%; Pred. No. 1.1e-129;  
Matches 268; Conservative 22; Mismatches 28; Indels 0; Gaps 0;

QY 8 IAKDVTLLIGKTPVLYLNKLDGCVARVAAKLELMEPCSSVKDRIGYSMTADAEKGLIT 67  
DB 11 IAKDVTLLIGKTPVLYLNKLDGCVARVAAKLELMEPCSSVKDRIGYSMTADAEKGLIT 70  
QY 68 PKSVLIBPTSGNTGIGLAFAAARGYKLIITMPASMSLERRIILAFGAELVLTDPAG 127  
DB 71 PGVSVLIBPTSGNTGIGLAFAAARGYKLIITMPASMSMERRIILAFGAELVLTDPAG 130  
QY 128 MKGAVOKAEIILAKTPNAYILQGFENPANKVHYETTGPEIWKSGDKIDAFVSGIGTGG 187  
DB 131 MKGAVOKAEIILAKTPNAYILQGFENPANKVHYETTGPEIWKATAGKIDAFVSGIGTGG 190  
QY 188 TITGAGKYLKEONPNKILIGVEPVESVLSGKPGPHKIQIGAGFIPGVLEVNLLDEVV 247  
DB 191 TITGGRYLRQNPVNLKYGVEPVESAVLNGKPGPHKIQIGAGFIPGVLDVLLDETL 250  
QY 248 QISSDEALETAKLALKEGLFVGISGAAAAAFQIAKRPENAGKLIIVAVPSPGGRYLS 307  
DB 251 QVSSDEALETAKLALKEGLLVGISSGAAAAAVRLAKRPENAGKLIIVAVPSPGGRYLS 310

QY 308 SVLFESVREAEASMTFEP 325  
 DB 311 SVLFQSIKKEAESMTVEP 328

RESULT 10

US-10-464-811-21  
 : Sequence 21, Application US/10464811  
 : Publication No. US2004003219A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Kraus, Jan  
 : APPLICANT: Oliveriusova, Jana  
 : TITLE OF INVENTION: Human Cystathionine B-Synthase Variants and Methods of Production  
 : FILE REFERENCE: 2848-49  
 : CURRENT APPLICATION NUMBER: US/10/464,811  
 : PRIOR FILING DATE: 2003-06-17  
 : PRIOR APPLICATION NUMBER: 60/389,541  
 : PRIOR FILING DATE: 2002-06-17  
 : NUMBER OF SEQ ID NOS: 22  
 : SOFTWARE: PatentIn version 3.1  
 : SEQ ID NO 21  
 : LENGTH: 325  
 : TYPE: PRT  
 : ORGANISM: *Triticum aestivum*  
 : US-10-464-811-21

Query Match 84.0%; Score 1364; DB 12; Length 325;  
 Best Local Similarity 82.3%; Pred. No. 1.4e-128;  
 Matches 261; Conservative 27; Mismatches 29; Indels 0; Gaps 0;

QY 8 IAKVTEHIGKTPVLYLNKLADGCVARVAAKLELMEPCSSVKDRIGYSMTADAEKELT 67  
 DB 9 IAKVTEHIGKTPVLYLNKLADGCVARVAAKLELMEPCSSVKDRIGYSMTADAEKELT 68  
 QY 68 PKGSVLEIPTSNTGIGLAFMAAARGYKLIITMPASMSLERIILLAFGAEVLVLPKAG 127  
 DB 69 PKGSVLEIPTSNTGIGLAFMAAARGYKLIITMPASMSMERIILLAFGAEVLVLPKAG 128  
 QY 128 MKGAVQKAEELIAPNAYILQGFENPANKVHYETTGPBEIWKSGDKIDAFVSGIGTGG 187  
 DB 129 MKGAVQKAEELIAPNAYILQGFENPANKVHYETTGPBEIWKSGDKIDAFVSGIGTGG 188  
 QY 188 TTGAGKILKEONPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEVNLDDEV 247  
 DB 189 TTGAGKILKEONPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEVNLDDEV 248  
 QY 248 QISSDEALIEETAKLILKEGLFVGISGAAAAAFOIAKRPENAGKLIIVAFPSFGERYLS 307  
 DB 249 QISSDEALIEETAKLILKEGLFVGISGAAAAAFOIAKRPENAGKLIIVAFPSFGERYLS 308  
 QY 308 SVLFESVREAEASMTFEP 324  
 DB 309 SVLFQSIKKEAESMTVEP 325

RESULT 11

US-09-931-457A-64  
 : Sequence 64, Application US/09931457A  
 : Patent No. US20020157132A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Falco, S. Carl  
 : APPLICANT: Allen, Stephen M.  
 : TITLE OF INVENTION: Plant Amino Acid Biosynthetic Enzymes  
 : FILE REFERENCE: B8116 US CIP  
 : CURRENT APPLICATION NUMBER: US/09/931,457A  
 : PRIOR FILING DATE: 2002-02-22  
 : PRIOR APPLICATION NUMBER: 09/424,976  
 : PRIOR FILING DATE: 1999-12-02  
 : PRIOR APPLICATION NUMBER: 60/065,385  
 : PRIOR FILING DATE: 1997-11-12  
 : PRIOR APPLICATION NUMBER: 60/049,406  
 : PRIOR FILING DATE: 1997-06-12

: NUMBER OF SEQ ID NOS: 72  
 : SOFTWARE: Microsoft Office 97  
 : SEQ ID NO 64  
 : LENGTH: 415  
 : TYPE: PRT  
 : ORGANISM: *Oryza sativa*  
 : US-09-931-457A-64

Query Match 81.1%; Score 1317; DB 9; Length 415;  
 Best Local Similarity 77.4%; Pred. No. 1.1e-123;  
 Matches 250; Conservative 39; Mismatches 34; Indels 0; Gaps 0;

QY 3 VERSGIADVTETLIGKTPVLYLNKLADGCVARVAAKLELMEPCSSVKDRIGYSMTADAE 62  
 DB 93 VEGINTADVTQLIGKTPVLYLNKLADGCVARVAAKLELMEPCSSVKDRIGYSMTADAE 152  
 QY 63 KGLITPKGSVLEIPTSNTGIGLAFMAAARGYKLIITMPASMSLERIILLAFGAEVLVLT 122  
 DB 153 KGLITPKGSVLEIPTSNTGIGLAFMAAARGYKLIITMPASMSMERIILLAFGAEVLVLT 212  
 QY 123 DPAKMGAVQKAEELIAPNAYILQGFENPANKVHYETTGPBEIWKSGDKIDAFVSG 182  
 DB 213 DPAKMGAVQKAEELIAPNAYILQGFENPANKVHYETTGPBEIWKSGDKIDAFVSG 272  
 QY 183 IGTGCTTGAGKILKEONPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEVNLT 242  
 DB 273 IGTGCTTGAGKILKEONPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPGVLEVNLT 332  
 QY 243 LDBVVOISSDEALIEETAKLILKEGLFVGISGAAAAAFOIAKRPENAGKLIIVAFPSFG 302  
 DB 333 LDBVVOISSDEALIEETAKLILKEGLFVGISGAAAAAFOIAKRPENAGKLIIVAFPSFG 392  
 QY 303 ERYLSSVLFESVREAEASMTFEP 325  
 DB 393 ERYLSSVLFESVREAEASMTFEP 415

RESULT 12

US-09-931-457A-66  
 : Sequence 66, Application US/09931457A  
 : Patent No. US20020157132A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Falco, S. Carl  
 : APPLICANT: Allen, Stephen M.  
 : TITLE OF INVENTION: Plant Amino Acid Biosynthetic Enzymes  
 : FILE REFERENCE: B8116 US CIP  
 : CURRENT APPLICATION NUMBER: US/09/931,457A  
 : PRIOR FILING DATE: 2002-02-22  
 : PRIOR APPLICATION NUMBER: 09/424,976  
 : PRIOR FILING DATE: 1999-12-02  
 : PRIOR APPLICATION NUMBER: 60/065,385  
 : PRIOR FILING DATE: 1997-11-12  
 : PRIOR APPLICATION NUMBER: 60/049,406  
 : PRIOR FILING DATE: 1997-06-12  
 : NUMBER OF SEQ ID NOS: 72  
 : SOFTWARE: Microsoft Office 97  
 : SEQ ID NO 66  
 : LENGTH: 386  
 : TYPE: PRT  
 : ORGANISM: *Solanum tuberosum*  
 : US-09-931-457A-66

Query Match 81.0%; Score 1314; DB 9; Length 386;  
 Best Local Similarity 78.1%; Pred. No. 1.9e-123;  
 Matches 249; Conservative 38; Mismatches 32; Indels 0; Gaps 0;

QY 3 VERSGIADVTETLIGKTPVLYLNKLADGCVARVAAKLELMEPCSSVKDRIGYSMTADAE 62  
 DB 64 VEGINTADVTQLIGKTPVLYLNKLADGCVARVAAKLELMEPCSSVKDRIGYSMTADAE 123  
 QY 63 KGLITPKGSVLEIPTSNTGIGLAFMAAARGYKLIITMPASMSLERIILLAFGAEVLVLT 122  
 DB 124 KGLITPKGSVLEIPTSNTGIGLAFMAAARGYKLIITMPASMSMERIILLAFGAEVLVLT 183



QY 123 DPAKMGKAVQKAEELIAPNAYILQOFENPANPKVHYETTGTPEIWKSGSDGKIDAFVSG 182  
 DB 184 DPAKMGKAVQKAEELIAPNAYILQOFENPANPKVHYETTGTPEIWKSGSDGKIDAFVSG 243  
 QY 183 IGTGCTITGAGKYLKEONPNKILIGVEPVSFVLSGKPGPHKIQIGAGFIPGVLEVN 242  
 DB 244 IGTGCTITGAGKYLKEONPNKILIGVEPVSFVLSGKPGPHKIQIGAGFIPGVLEVN 303  
 QY 243 LDEVVQISSDEAIEFTAKLALKEGLFVGISGAAAAAFAQIAKRPENAGKLIIVAFPSFG 302  
 DB 304 MDEVIEISSDEAIEFTAKLALKEGLFVGISGAAAAAFAQIAKRPENAGKLIIVAFPSFG 363  
 QY 303 ERYLSSVLPESVREAESM 321  
 DB 364 ERYLSSVLPESVREAESM 382

RESULT 13

US-10-425-114-60706  
 ; Sequence 60706, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jindong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 60706  
 ; LENGTH: 316  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: LIB3587-229-H7\_FLI.pep  
 ; US-10-425-114-60706

Query Match 80.7%; Score 1309; DB 12; Length 316;  
 Best Local Similarity 81.1%; Pred. No. 4.5e-123;  
 Matches 258; Conservative 22; Mismatches 26; Indels 12; Gaps 1;

QY 8 IAKVTEELIGKTPLVYLNKLADGCVARVAAKLEIMBPCCSVKDRIGYSMIADAEKGLIT 67  
 DB 11 IAKVTEELIGKTPLVYLNKLADGCVARVAAKLEIMBPCCSVKDRIGYSMIADAEKGLIT 58  
 QY 68 PGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLEERILLAFGAELVLTDPAG 127  
 DB 59 PGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLEERILLAFGAELVLTDPAG 118  
 QY 128 MKGAVQKAEELIAPNAYILQOFENPANPKVHYETTGTPEIWKSGSDGKIDAFVSGITG 187  
 DB 119 MKGAVQKAEELIAPNAYILQOFENPANPKVHYETTGTPEIWKSGSDGKIDAFVSGITG 178  
 QY 188 TITGAGKYLKEONPNKILIGVEPVSFVLSGKPGPHKIQIGAGFIPGVLEVNLDDEV 247  
 DB 179 TITGAGKYLKEONPNKILIGVEPVSFVLSGKPGPHKIQIGAGFIPGVLEVNLDDEV 238  
 QY 248 QISSDEAIEFTAKLALKEGLFVGISGAAAAAFAQIAKRPENAGKLIIVAFPSFGERYLS 307  
 DB 239 QISSDEAIEFTAKLALKEGLFVGISGAAAAAFAQIAKRPENAGKLIIVAFPSFGERYLS 298  
 QY 308 SVLPESVREAESMTFEP 325  
 DB 299 SVLPESVREAESMTFEP 316

RESULT 14

US-10-425-114-57303  
 ; Sequence 57303, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jindong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 57303  
 ; LENGTH: 397  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: UC-ZMFLMO17258D02\_FLI.pep  
 ; US-10-425-114-57303

Query Match 78.8%; Score 1279; DB 12; Length 397;  
 Best Local Similarity 74.9%; Pred. No. 6.7e-120;  
 Matches 245; Conservative 44; Mismatches 36; Indels 2; Gaps 1;

QY 1 MAVERSG-IADVTEELIGKTPLVYLNKLADGCVARVAAKLEIMBPCCSVKDRIGYSMTA 58  
 DB 71 VAAEVGGLINADVTLQIGNTPMVYLNKVVGVANVAKLEIMBPCCSVKDRIGYSMTA 130  
 QY 59 DAEKGLITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLEERILLAFGA 118  
 DB 131 DAEKGLITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLEERILLAFGA 190  
 QY 119 LVLTPAKMGKAVQKAEELIAPNAYILQOFENPANPKVHYETTGTPEIWKSGSDGKIDA 178  
 DB 191 LVLTPAKMGKAVQKAEELIAPNAYILQOFENPANPKVHYETTGTPEIWKSGSDGKIDA 250  
 QY 179 FVSGIGTGTITGAGKYLKEONPNKILIGVEPVSFVLSGKPGPHKIQIGAGFIPGV 238  
 DB 251 FVSGIGTGTITGAGKYLKEONPNKILIGVEPVSFVLSGKPGPHKIQIGAGFIPGV 310  
 QY 239 EVNLDDEVVQISSDEAIEFTAKLALKEGLFVGISGAAAAAFAQIAKRPENAGKLIIVAF 298  
 DB 311 EVNLDDEVVQISSDEAIEFTAKLALKEGLFVGISGAAAAAFAQIAKRPENAGKLIIVAF 370  
 QY 299 PSFGERYLSVLPESVREAESMTFEP 325  
 DB 371 PSFGERYLSVLPESVREAESMTFEP 397

RESULT 15

US-10-425-114-71564  
 ; Sequence 71564, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jindong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 71564  
 ; LENGTH: 398  
 ; TYPE: PRT

ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: UC-ZMFLMO17281G01\_FLI.pep  
US-10-425-114-71564

Query Match 78.8%; Score 1279; DB 12; Length 398;  
Best Local Similarity 74.9%; Pred. No. 6.7e-120;  
Matches 245; Conservative 44; Mismatches 36; Indels 2; Gaps 1;

QY 1 MAVERSG--IAKDVTELLIGKTPLVYLNLKLDGCVARVAAKLELMEPCSSVKDRIGYSMTA 58  
DB 72 VAAEVGGILINANDVTQIGNTPMVYLNINNVKGSVANVAAKLEIMEPCSSVKDRIGYSMTA 131  
QY 59 DAEKGLITPGKSVLIEPTSGNTGIGLAFMAARGYKLIITMPASMSLERRIILAFGAE 118  
DB 132 DAEKGLITPGKSVLIEPTSGNTGIGLAFMAARGYKLIITMPASMSLERRIILAFGAE 191  
QY 119 LVLTDPKMGKAVQKAEELAKTPNAYILQGFENPANPKVHYETTPPEIWKSGSDGKIDA 178  
DB 192 LVLTDPKMGKALDKATEILNKTPNSYMLQGFENPANPKVHYETTPPEIWKSGSDGKIDA 251  
QY 179 FVSGIGTGTITGAGKYLKQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPGV 238  
DB 252 FIGIGTGTITGAGKYLKQNPNIKLIGVEPVESPVLSGKPGPHKIQIGAGFIPGV 311  
QY 239 EVNLDEVVQISSDEALETAKLALKEGLFVGISSGAAAAAFQIAKRPENAGKLIYAVF 298  
DB 312 DSDILDEVIEISSDEAVETAKQAVQEGILVGISSGAAAAAFQIAKRPENAGKLIYAVF 371  
QY 299 PSFGERYLSVLFESVREARESMTEP 325  
DB 372 PSFGERYLSVLYQSTIRECEENMQPEP 398

Search completed: May 18, 2004, 13:14:30  
Job time : 49 secs